A Probabilistic Approach to some Problems of Number Theory

Gregory M. Sobko

Abstract. Some classical questions and problems of Number Theory, like the Goldbach conjecture [1], distributions of twin- and d-primes and primes among arithmetic sequences, are addressed here from an entirely probabilistic point of view. We discuss the concepts of 'randomness' and 'independence' relevant to number-theoretic problems and interpret the basic concepts of divisibility of natural number in terms of probability spaces and appropriate probability distributions on classes of congruence. We analyze and demonstrate the importance of Zeta probability distribution and prove, in particular, theorems stating the equivalence of probabilistic independence of divisibility by coprime factors, and the fact that random variables with the property of independence of coprime factors must have Zeta probability distribution.

Multiplicative and additive models with recurrent equations for generating sequences of prime numbers are derived based on the reduced Sieve of Eratosthenes Algorithm. This allows to interpret such sequences as realizations of random walks on set $\mathbb{N}$ of natural numbers and on multiplicative semigroups $S(\mathbb{P})$ generated by set of prime numbers $\mathbb{P}$, representing paths of stochastic dynamical systems. The H. Cramér model for probability distribution of primes is modified as a *generalized predictable non-stationary Bernoulli process with dependent terms,* which are *asymptotically pairwise Bernoullian*, and applied to analyze the sequences of primes generated by appropriate random walks. This leads to study of prime numbers distributions among arithmetic sequences (classes of congruence) based on expectations and variances for occurrences of primes not exceeding x in arithmetic sequences. We illustrate this by computer calculations supporting the

conjecture of the uniform distribution of primes among congruence classes for each given prime p. With an intense use of Zeta probability distribution it seems possible by using the modified Cramér's model to prove the Goldbach conjecture. The solution to the Twin Primes problem and more general d-primes distribution problem for consecutive prime numbers is also suggested in this paper. We discuss some limit theorems related to distribution of primes and their residuals. More specifically, we provide a continuous-time description of the distribution of counting function of primes $\pi(n)$, $n \in \mathbb{N}$, in terms of diffusion approximation of non-Markov random walks.

PREFACE

"*Any great quest demands courage. It is a voyage into unknown with no guaranteed results…Our lives also have this quality of a quest, the attempt to resolve some fundamental but ill-posed question. In working on a mathematical conjecture, life's ambiguities solidify into a concrete problem…*
This is one reason that working on mathematics is so satisfying.
In resolving the mathematical problem, we for a while at least, resolve that large, existential problem that is consciously or unconsciously always with us…"
(William Byers*, How Mathematicians Think,* Princeton University Press, 2007)

"…Mathematics is the art of giving the same name to different things…The only facts worthy of our attention are those which introduce order into this complexity and so make it accessible to us".
(Henry Poincaré, The Value of Science, Random House, Inc., 2001.

"Using randomness to study certainty may seem somewhat surprising. It is, however, one of the deepest contributions of our century to mathematics in general and to the theory of numbers in particular." (Gérald Tenenbaum, Michel Mendès France, *The Prime Numbers and Their Distribution*. AMS, 2000)

Number Theory is a precious and an inexhaustible source of problems, ideas and methods that inspire the development of mathematical sciences starting from the dawn of human civilization. This work is dedicated to some classical questions and problems of Number Theory that I have attempted to address from an entirely probabilistic point of view by using quite elementary methods. To approach Number Theory issues we have to discuss existence and representation of probability spaces and probability distributions relevant to divisibility problems, prime factorization, distribution of prime numbers, the Prime Number Theorem. The work is focused on a special role of Riemann Zeta probability distribution (associated with Zeta function) on the set $\mathbb{N}$ of natural numbers and on some modifications of the famous Cramér's model of probability distribution of primes [2,9].
A detailed analysis of Riemann zeta distribution and its connections to Number Theory are presented by Gwo Dong Lin and Chin-Yuan Hu in [11] and by many other authors. This work has a number of inevitable intersections with other studies, though I tried to avoid common places.
We discuss the divisibility problem of a 'random number' ν in terms of statistical independence to show that the statement of Mark Kac [4, p. 53] that "primes play a game of chance" does not carry only a metaphorical meaning,

but also has a strict mathematical sense, assuming that 'random number' ν follows a specific probability distribution.

With certain intense use of Zeta probability distribution, we approached some of old classical problems in Number Theory like the strong form of Goldbach problem and the Twin Primes problem (the last generalized to the d-primes distribution problem for consecutive prime numbers), and the distribution of prime numbers among arithmetic sequences.

Sequences of natural numbers in this work are considered as realizations of paths of multiplicative random walks with independent increments on $\mathbb{N}$ (generated by random variables ν followed Zeta distribution), while prime-valued sequences are represented as realizations of additive random walks with dependent increments on $\mathbb{N}$. We denote here $\mathbb{P}$ the set of prime numbers.

Better foundations for the Cramer's model in this work is provided by considering the sequence $\left(\xi(\nu_n)\right)_{n\in\mathbb{N}}$ of $(0,1)$-valued random variables $\xi(\nu_n)=\begin{cases}1 \text{ if } \nu_n\in\mathbb{P}\\ 0 \text{ otherwise}\end{cases}$ as a *generalized predictable non-stationary Bernoulli process with dependent terms*. As it is shown in Theorem 3.1, the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ in the modified Cramér's model is *asymptotically pairwise Bernoullian,* that is $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|=O\left(\dfrac{1}{\ln N}\right)$, where $P\{\xi_k=1\}=P_k$, $P\{\xi_k\cdot\xi_l=1\}=P_{kl}$, $X_n=\sum\limits_{k=1}^{n}\xi_k$, and $\left|V\{X_n\}-\dfrac{1}{n^2}\sum\limits_{k=1}^{n}V\{\xi_k\}\right|=O\left(\dfrac{1}{\ln N}\right)$.

The sequence of random variables $\left(\nu_n\right)_{n\in\mathbb{N}}$ is such that $\nu_n(\omega)=p_n\in\mathbb{P}$ is a realization of the random sequence $\left(\nu_n\right)_{n\in\mathbb{N}}$, representing primes for some $\omega\in\Omega$ in the main probability space $(\Omega,\mathcal{F},P)$. The assignments of probabilities

$$P\{\nu_n = p_n\} = P\{\xi(n)=1\} = \frac{1}{\ln n},\ P\{\nu_n \neq p_n\} = P\{\xi(n)=0\} = 1-\frac{1}{\ln n},$$

in the Cramér's model was originally motivated by the Prime Number Theorem [10, p.133], where the counting function of primes on $\mathbb{N}$ is given by the asymptotic formula $\pi(x) = \sum_{p\in\mathbb{P}\cap[2,x]} 1 \sim Li(x) = \int_2^x \frac{dt}{\ln t}$, which leads to the heuristic assumption about the probability $P\{p\in[x-1,x]\} \sim \int_{x-1}^x \frac{dt}{\ln t} \sim \frac{1}{\ln x}$.

By assuming that each random variable ν_n in the sequence $(\nu_n)_{n\in\mathbb{N}}$ follows Riemann Zeta probability distribution $P_s\{\nu = n\} = \frac{n^{-s}}{\zeta(s)},\ n\in\mathbb{N}, s>1$, where $\zeta(s) = \sum_{n\in\mathbb{N}} n^{-s}, s>1$, is Riemann zeta function, we prove that $P_s\{\nu\in\mathbb{P}|\nu = n\} = \prod_{p\le\sqrt{n}}\left(1-\frac{1}{p^s}\right)$ (Theorem 2.2) . Then, as it follows from the Merten's 2nd theorem [, pp. 18-19], we have the asymptotic expression:

$$P\{\xi(n)=1\} = \prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right) = \frac{e^{-\gamma}}{\frac{1}{2}\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right] = \frac{c}{\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right],$$

where $c = \frac{2}{e^\gamma} \approx 1.122918968$, γ denotes Euler's constant.

Given that a random variable ν follows Zeta probability distribution $P_s\{\nu = n\} = \frac{n^{-s}}{\zeta(s)}, s>1, n\in\mathbb{N}$, we discuss the following issues:

1) probability distribution of exponents $\alpha(\nu,p)$ in the prime factorization $\nu = \prod_{p\in\mathbb{P}} p^{\alpha(\nu,p)}$;

2) a multiplicative model of random walks $\nu(k+1)=\nu(k)\cdot\xi(k+1)$,where $\nu(0)=1$, $\xi(k)=p_k^{\alpha(\nu(k),p_k)}$ $(k=1,2,\ldots)$ on semigroups $S(\mathbb{P})$ and $S(\mathbb{P}_N)$ generated by

all primes $\mathbb{P}$ and by a limited set of primes $\mathbb{P}_N=\{p\in\mathbb{P}\mid p\le N\}$, respectively.

3) probabilistic interpretation of Riemann Zeta function $\zeta(z)=\sum_{n\in\mathbb{N}} n^{-z}$ in the derived formula $\zeta(z)=f_{\lambda_s}(t)\cdot\zeta(s)$, where $z=s+i\cdot t$ is in the domain of convergence of $\zeta(z)$ and $f_{\lambda_s}(t)=E\{e^{i\cdot t\cdot\lambda_s}\}$ is a characteristic function of a random variable λ_s $(s>1)$ with probability distribution

$P_s\{\lambda_s=-\ln n\}=\dfrac{1}{n^s\cdot\zeta(s)}$ $(n\in\mathbb{N})$;

4) independence of divisibility of ν by coprime factors as a necessary and sufficient condition for probability distribution of ν to be a Zeta Riemann distribution;

5) improvement of Cramér's model represented as a non-stationary asymptotically pairwise Bernoulli process on the reduced Eratosthenes Sieve algorithm and Mertens theorems to approximate prime distribution function $\dfrac{\pi(x)}{x}$ as $x\to\infty$.

6) By using the Dirichlet characters for a finite abelian group $G_p=\mathbb{Z}_p=\mathbb{Z}/(p\cdot\mathbb{Z})$, $p\in\mathbb{P}$, and the corresponding characteristic functions, we discuss distribution of primes among arithmetic sequences and asymptotic distribution of residuals $r=\mathrm{mod}(\nu,p)=[\nu]_p$, $p\in\mathbb{P}$. We prove (Theorem 4.2) that for a sequence $\nu_1,\nu_2,\ldots,\nu_n,\ldots$ of independent random integers (not necessarily equally distributed), the residuals of sums

$\left[v^{(n)}\right]_p = \sum_{i=1}^{n} [v_i]_p$ are asymptotically uniformly distributed on G_p, for every $p \in \mathbb{P}$, and components of the vector of residuals $\vec{r}(v) = \left(r_1, r_2, \ldots r_{\pi(v)}\right)$ are asymptotically independent random variables.

Then, for arithmetic sequences $C_{pr} = \left\{n = k \cdot p + r \mid k \in \mathbb{N}\right\}$, $1 \le r \le p-1$, and set $\mathbb{P}_{pr} = C_{pr} \cap \mathbb{P} \setminus \{p\}$ of primes in these sequences (Theorem 5.1) we show that $\left[v_n\right]_p \to v_0$ in probability, where random variable v_0 has a uniform distribution on $\{01, 2, \ldots, p-1\}$.

7) Proving that every $D\mathbb{P}_d$ is an infinte set for all even values of $d \ge 2$), where $D\mathbb{P}_d = \{p \mid p \text{ and } p+d \text{ are consecutive primes}\}$ is the set of d-primes (that is prime numbers p such that p and $p+d$ are consecutive primes). Notice that $D\mathbb{P}_2$ is the set of twin-primes. Assuming the Cramer's assumption of independence of consecutive primes, supported by the fact that the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ in the modified Cramér's model is *asymptotically pairwise Bernoullian,* we prove that every $D\mathbb{P}_d$ is an infinte set for all even values of $d \ge 2$) .
8) By using similar arguments, we approach the strong Goldbach conjecture and try to 'solve' the puzzle in the framework of probability theory, by using the modified (see above) H. Cramér's assumption of independence of primes occurred in the sequence of natural numbers $\mathbb{N}$. We consider the so-called *Goldbach function* $G(2m)$ that denotes the number of presentations of an even integer in the form: $2m = p + p'$ where p, p' are prime numbers (called G-primes). A choice of a G-prime p for

every $m \geq 3$ is considered as a realization of $G(2m,v)$ for a random variable v with Zeta probability distribution. We have then, $2m = v + v'$ where $v \in \mathbb{P},\ v' \in \mathbb{P}$. The calculations for the available range of v values show that the number of representations $G(2m)$ of an even integer in the form $2m = p + p'$ where p, p' are primes, increases when m increases and becomes larger for the larger values of m. A prime number $p \in \mathbb{P},\ p \leq m$, we call a G_m- *prime* if there exist an even number $2m \geq 6$ and a prime number $p' \in \mathbb{P}$ such that $2m = p + p'$. The set of all G_m-*primes* for a given m we denote $G_m\mathbb{P}$. The main results of this section are stated in Theorems 7.1 and 7.2. The most critical question for the Goldbach strong conjecture is whether the probability that for 'sufficiently large' values of $m > M \geq 3$ all sets $G_m\mathbb{P}$ are not empty, or equivalently, is this true that

$P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,v_m) = |G_m\mathbb{P}| > 0\right\}\right\} \to 1 \text{ as } M \to \infty$. Let $G_m\mathbb{P}$ for $m \geq 3$ be a set of all G-primes, that is prime numbers $p, p' \in \mathbb{P}$ such that $p + p' = 2m$. Assume that each random variable v_m in the sequence of independent random variables $\{v_m\}_{3\leq m}$ follows Zeta probability distribution:

$$P\{v_m = n\} = \frac{n^{-s}}{\zeta(s)} \quad (s > 1) \text{ and } \gamma_m(v_m) = \begin{cases} 1 \text{ if } v_m = n \in G_m\mathbb{P} \\ 0 \text{ otherwise} \end{cases}.$$

Then, $\{\gamma_m(v_m)\}_{m\geq 3}$ as a sequence of asymptotically independent Bernoulli variables implies the following properties of Goldbach function.

$G(2m,v_m) = \sum_{n=3}^{2m-3}\gamma_m(v_m)$:

(1) $$P\{G(2m,v_m) = 0\} = P\left\{\bigcap_{n=3}^{2m-3}\{\gamma_m(v_m) = 0 \mid v_m = n\}\right\} \to 0 \text{ as } m \to \infty.$$

(2) $\sum_{m=3}^{\infty} P\left\{G(2m,v_m)=0\right\} < \sum_{m=3}^{\infty} e^{-\frac{2m-6}{\ln^2(2m)}} \approx 6.00236$

(3) $\lim_{M\to\infty} P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,v_m)=|G_m\mathbb{P}| > 0\right\}\right\} \to 1$

Due to Lemma 3.2, the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ in the Cramér's model is *asymptotically pairwise Bernoullian,* so that

$$\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\frac{1}{\ln N}\right) \text{ as } D_n=O\left(\frac{1}{\ln N}\right) \text{ for all } n>N.$$

The last property implies that in the formula $G(2m,v)=\sum_{n=3}^{2m-3}\gamma_m(n)>0$ we have all random variables $\gamma_m(n)$, $\mathrm{N}\le n\le 2m-3$, not necessarily independent but asymptotically linearly uncorrelated. Then, $\lim_{m\to\infty} P\left\{G_m\mathbb{P}\ge 1\right\}=1$.

9) Given a sequence $\left(\xi(n)\,|\,n\in\mathbb{N}\right)$ of prime numbers indicators as a non-stationary time series, we evaluate covariances (and correlations) between $\xi(k)$ and $\xi(k')$ for arbitrary k and k', and consider two situations in the study of correlations between $\xi(k)$ and $\xi(k')$: (1) k and k' are close enough to each other; (2) k and k' are separated from each other by 'long' intervals. In both cases computer calculations show the low values of range for the sample correlation coefficients:

Range of a sample correlation coefficients $R=\left[\mathrm{co}r\left(\vec{\xi}_i,\vec{\xi}_{i+1}\right)\right]$ $(i=1,2,...,n-1)$ for $n=100$ consecutive $m=10^4$-intervals of allocation of primes $< 10^6$: $0.03776633\le R \le 0.09711712$.

10) Finally, a continuous-time description of the distribution of counting function of primes $\pi(n),\ n\in\mathbb{N}$, is obtained in terms of diffusion approximation of non-Markov random walks. In this chapter we consider

the sequence $\{\pi(n)\}_{n\in\mathbb{N}}$ as a realization of a random walks $\{\pi(n,\omega)\}_{n\in\mathbb{W}}$ generated by the recurrent equation $\pi(n_{k+1})-\pi(n_k)=\eta(n_{k+1})$ where $\eta(n_k)=h\left(\min(\vec{r}(n_k)\right)$, $n_k=\nu_k(\omega)$. Here $\{\nu_k\}_{k\in\mathbb{N}}$ are assumed to be random variables with Zeta probability distribution.

We consider a stochastic process approximation of non-Markov random walks $\{\pi(n,\omega)\}_{n\in\mathbb{W}}$ such that $\pi(n,\omega)=\pi(n)$, with $\pi:\mathbb{N}\times\Omega\to\mathbb{N}\cup\{0\}$ restricted to the interval of discrete 'times' $N_{\min}=n_0<n_1<\cdots<n_K=N_{\max}$:

$$\left\{\pi^{\Delta}(t_k)=\pi(n_k,\omega)\mid N_{\min}\le n_k\le N_{\max}\right\}.$$

We denote $\Delta=\left(0=t_0<t_1<\ldots<t_K=1\right)$ a partition of an interval $[0,1]$ into K subintervals, such that $\frac{K}{\ln\left(N_{\max}\right)}\to 0\ as\ N_{\max}\to\infty$.

The closed interval of real numbers $\left[N_{\min},N_{\max}\right]\subset\mathbb{R}$ is mapped to the interval $[0,1]\subset\mathbb{R}$ by an increasing continuously differentiable function $\tau(x)$ such that $\tau(N_{\min})=0,\ \tau(N_{\max})=1$ where $\tau(x)=\frac{\int_{N_{\min}}^{x}\frac{dt}{\ln t}}{\int_{N_{\min}}^{N_{\max}}\frac{dt}{\ln t}}=\frac{Li(x)-Li(N_{\min})}{Li(N_{\max})-Li(N_{\min})}$

and $Li(x)$ stands for the *Eulerian logarithmic integral* $Li(x)=\int_2^x\frac{dt}{\ln t}$.

Then, $t_k=\tau(n_k)$ and for τ^{-1} (the inverse of τ) we have $n_k=\tau^{-1}(t_k)$ $(k=1,2,\ldots,K)$. Denote $\Delta t_k=t_k-t_{k-1}$ and assume that $N_{\min}\to\infty$ and for each choice of $N_{\min}$ a positive integer K can be taken such that $|\Delta|=\max_{1\le k\le K}\Delta t_k\to 0$.

Here a sequence of random variables $\pi^{\Delta}(t_k)=\pi(n_k)$ is interpreted as a path of a walking point $\pi^{\Delta}(t_k)$ that belongs to a measurable space $\left(\mathcal{X}_k,\mathcal{B}_k\right)$ at each 'instant of registration' t_k. Theorem 9.1 states that

transition probabilities $P\left\{\pi^{\Delta}(t_{k+1}) \in E \mid \pi^{\Delta}(t_k) = x_k, \pi^{\Delta}(t_{k-1}) = x_{k-1}, \ldots, \pi^{\Delta}(t_0) = x_0\right\}$,

where $\vec{x}_k = (x_1, \ldots, x_k) \in \mathbb{N}^k$, of the defined above non-Markov sequence of random walks *converges weakly* to the transition probabilities of the diffusion process $\hat{\pi}(t)$ given by the stochastic integral

$$\hat{\pi}(t) = \int_0^t \hat{m}(s)\, ds + \int_0^t \hat{\sigma}(s)\, dw(s),$$

where $\hat{m}(t) = \dfrac{c}{\ln\left(\tau^{-1}(t)\right)}, \hat{\sigma}(t) = \dfrac{1}{2} \cdot \hat{m}(t) \cdot \left(1 - \hat{m}(t)\right), c = \dfrac{2}{e^{\gamma}}, 0 \le t \le 1,$

$\tau^{-1}(t) = x,\ N_{\min} \le n \le N_{\max}, \tau(N_{\min}) = 0, \tau(N_{\max}) = 1,\ c = \dfrac{2}{e^{\gamma}} \approx 1.122918968$

with the Euler's constant $\gamma = \sum_{m \le n} \dfrac{1}{m} - \ln n + O\left(\dfrac{1}{n}\right), \gamma \approx 0.577215664$,

as $\left|\Delta\right| = \max_{1 \le k \le K} \Delta t_k \to 0,\ N_{\min} \to \infty$.

The sequence of vectors $\left(\vec{p}(n), \vec{r}(n)\right), \left(n = 2, 3, \ldots\right)$ created by consecutive n primes and the residual values $\vec{r} = \mathrm{mod}(n, \vec{p})$, allows an interesting 3D presentation. In each pair $\left(\vec{p}(n), \vec{r}(n)\right)$ vector of primes $\vec{p}(n)$ represents a 'radial' component, while the vector of residuals $\vec{r}(n)$, due to its natural periodicity, represents a 'circular' component. As a result, we represent the sequence of consecutive primes numbers with the corresponding residual values $z_k = p_k \cdot \exp\left(2\pi i \cdot \dfrac{r_k}{p_k - 1}\right), r_k = \mathrm{mod}(n, p_k)$, $(k = 1, 2, 3, \ldots)$ on the complex plane C as a 3D helix.

The key issue in the probabilistic analysis in a number-theoretic framework remains an enigmatic connection between deterministic nature of integer sequences related to prime numbers and their apparent complicated (‘unpredictable’ or ‘chaotic’) behavior interpreted as ‘randomness’.

The concept of ‘randomness’ is in the focus of numerous philosophical and mathematical studies and discussion. It follows different interpretations among physicists, engineers, biologists, specialists in information technology, computer scientists, and of course, mathematicians, as outlined by Edward Beltrami in [6 , p. 92].

An apparently ‘random’ patterns of some sequence of integers are naturally related to their ‘complexity’ or ‘predictability’, which leads to the concept of ‘algorithmic randomness’.

Due to this concept, complexity of a binary string can be measured by the length of an algorithm (written itself in as a binary string), which can generate the given binary string. To reproduce a string of maximum complexity (interpreted as ‘algorithmically random’), we need an algorithm which length is comparable with the length of the given string. Given a natural number n, one can use the Eratosthenes algorithm to calculate the count $\pi(n)$ of all consecutive prime numbers $p_1, p_2, \ldots, p_{\pi(n)}$ not exceeding n. To find out whether n is a prime number would require not more than $\pi(\sqrt{n})$ division operations for all primes $p \le \sqrt{n}$. This means that the sequence of consecutive primes has a ‘recursive memory’ of the size $\pi(\sqrt{n})$, and a ‘recursive complexity’ of a prime sequence $p_1, p_2, \ldots, p_{\pi(n)}$ can be estimated, due to the Prime Number Theorem [12, 22], as $\pi\left(\sqrt{n}\right) \sim \dfrac{2\sqrt{n}}{\ln n}$.

Quite deterministic nature of prime numbers, due to the complexity of the generating algorithm, is mimicking ‘randomness’, and allows apply some of probabilistic instruments to analyze number-theoretic problems.

It was hardly possible to make this text self-contained and I presume that a reader would have prerequisites based on university courses of Algebra, Calculus, Mathematical Analysis and Probability Theory, and (what is the most important) the curiosity and interest in both Number Theory and Probability.

Notice that all computer calculation results demonstrated in this work are performed with R and Matlab scripts created by the author, and I take full responsibility for any possible inaccuracies or mistakes. I tried to follow the style which combines traditional theoretical and computational approaches as stated in [22]: "Discussion – Definition – Theorem – Algorithm – Example".

Gregory M. Sobko

San Diego, California

2015 - 2020

CONTENTS

Chapter **Page**

1 . Stochastic Predictable Sequences, Prime Numbers and Riemann Zeta Probability Distribution

Let $\mathbb{N}$ denote the set of natural numbers and $\mathbb{P}$ the set of all primes. Our major assumption follows the amazing Cramér's idea [9] to represent a sequence of prime numbers as a realization of a random sequence of integers $(v_n \mid n \in \mathbb{N})$ with an appropriate choice of their probability distributions. Pursuing this idea we address two problems:

1) the choice of an adequate probability distributions P_n for each v_n;

2) stochastic relationship among all v_n in the sequence.

We need several definitions [7].

Definition 1.1

Let $\{v_n \mid n \in \mathbb{N}\}$ be random variables defined on probablity space $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}_n = \sigma\{v_k \mid 1 \le k \le n\}$ a σ-algebra generated by $\{v_k \mid 1 \le k \le n\}$. We have: $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \cdots \subseteq \mathcal{F}$, and for each $n \in \mathbb{N}$, random variable v_n is $\mathcal{F}_n$-measurable. Then, the sequence $(v_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is called a *stochastic sequence*.

A stochastic sequence $(v_n, \mathcal{F}_n)_{n\in\mathbb{N}}$ is called *predictable* if for each $n \in \mathbb{N}$ there exists $k = k(n) < n$ such that v_n is $\mathcal{F}_{k(n)}$-measurable. A pedictable sequence we can write as $\left(v_n, \mathcal{F}_{k(n)}\right)_{n\in\mathbb{N}}$.

Predictability of a stochastic sequence $(v_n, \mathcal{F}_n)_{n\in\mathbb{N}} = \left(v_n, \mathcal{F}_{k(n)}\right)_{n\in\mathbb{N}}$ means that for each $n \in \mathbb{N}$ the probability distribution P_n of v_n is compeletly determined by (depends on) values taken by some (or all) veriables $\nu_1,\nu_2,\ldots,\nu_{k(n)}$, where $k(n)<n$. That is, in terms of conditional probabilities,

$$P_n\left\{\nu\in A\,|\,\mathcal{F}_{n-1}\right\}=P_n\left\{\nu\in A\,|\,\mathcal{F}_{k(n)}\right\} \text{ for all } A\in\mathcal{F}_n\ . \tag{1.1}$$

Notice that general stochastic sequences include classes of sequences of independent as well as dependent random variables like martingales, Markov chains, etc.

In Number Theory we are interested in recursively defined sequences of numbers, generated by certain recurrent relations, mostly nonlinear. From probabilistic point of view, such recurrent relations generate sequences of dependent random variables. The problem of dependence of events and random variables in the framework of Number Theory had been dicussed in some detail in the monograph of Mark Kac [4]. As M. Kac underlined in [4], the concept of independence "though of central importance in probability theory, is not a purely mathematical notion", and it appears quite naturally in Statistical Physics. He mentioned that "the rule of multiplication of probabilities of independent events is an attempt to formalize this notion and to build a calculus arount it". By using informal language, the concept of independence is stated in [14] as follows: "Two events are said to be independent if they have 'nothing to do' with each other". To decide whether a 'randomly choosen' (odd) integer $\nu>2$ is a prime number, we subject ν to a divisibility test, according to the Eratosthenes algorithm. Then, the events $A=\left\{p_i\backslash\nu\right\}$ (' p_i divides ν ') and $B=\left\{p_j\backslash\nu\right\}$ (' p_j divides ν ') for $i\neq j$ do not depend on each other logically or statistically, and should be considered as independent for a 'reasonable' choice of probability distribution of random variable ν . As we demonstrate below, such a choice is provided by Zeta probability distribution

$$P\left\{v_m = n\right\} = \frac{n^{-s}}{\zeta(s)} \quad (s>1), \quad n \in \mathbb{N} \tag{1.2}$$

Both dependence and independence of 'events' in Number Theory are results of complicated recurrent nonlinear relations between terms of numeric sequences, which can generate a 'dynamical chaos', imitating pseudo-randomness in the long run behavior of such (theoretically) purely deterministic sequences. The precise prediction of behavior of terms in the sequences demands for 'big' numbers almost infeasible calculations caused by expanding memory of prehistory of their evolution. To make a study feasible and overcome "the curse of dependence", researchers in this area typically suggest heuristic assumptions that terms in a $\left(v_n\right)_{n\in\mathbb{N}}$ are independent, or asymptotically independent, or uncorrelated, or 'weakly' dependent, in a certain sense.

The basic fact is that the set of prime numbers $\mathbb{P}$ is a *recursive set* [18].

We can prove this by using an indicator function $I_{\mathbb{P}} : \mathbb{N} \to \left\{0,1\right\}$ of set $\mathbb{P}$. We need to show that the function $I_{\mathbb{P}}$ is recursively defined. (1) *Initial step*: let $I_{\mathbb{P}}(2) = 1$, $I_{\mathbb{P}}(3) = 1$. (2) *Inductive step*: if $n > 3$ is the smallest number such that $n \nmid k$ (symbol $\nmid$ means 'does not divide') for each $k \le \sqrt{n}$, then $I_{\mathbb{P}}(n) = 1$, otherwise $I_{\mathbb{P}}(n) = 0$. Notice that such number n exists. (3) *Closure step*: Only numbers n obtained in steps (1) and (2) satisfy condition $I_{\mathbb{P}}(n) = 1$.

It holds true that if a function is recursively defined then it is unique [18].

We can explain the above statement concerning the recursive definition of prime numbers as follows. Occurrence of a prime number $n = p \in \mathbb{P}$ in the sequence of consecutive natural numbers $n = \left\{2,3,4,\ldots\right\}$ depends on the values of reminders $r = \operatorname{mod}(n, p)$ for all primes $p \le n$, due to the Sieve of Eratosphenes Algorithm [5]. Actually, this requirement can be relaxed: we need to consider only divisibility of

n by all primes $p \le \sqrt{n}$. The proof of this statement (attributed to Fibonacci) follows below.

Lemma 1.1

A natural number $n \ge 5$ is prime if and only if n is not divisible by of any prime numbers $p \le \sqrt{n}$, or, equivalently, if $r = \mathrm{mod}(n, p) \ne 0$ for all primes $p \le \sqrt{n}$.

Proof.

If we assume that n is a composite number with no primes $p \le \sqrt{n}$ that divide n, then n should be divided by primes p_1' and p_2' both greater than $\sqrt{n}$, and therefore also divided by their product $p_1' \cdot p_2'$. But this would imply that $p_1' \cdot p_2' > n$, which is impossible.

This means that if n is not divisible by any of prime numbers $p \le \sqrt{n}$, then n itself must be a prime number.

Q.E.D.

The above discussion implies that the sequence of consecutive primes can be considered as a realization of a predictable stochastic sequence $\left(v_n, \mathcal{F}_{k(n)}\right)_{n \in \mathbb{N}}$, where $k(n) = \sqrt{n}$ for all $n > 3$.

One of the most challenging problems of Number Theory is the distribution of primes in the set $\mathbb{N}$ of natural numbers. The sequence of consecutive prime numbers $(2,3,5,7,11,\ldots)$ may look like a path of sporadic walks $\omega : \mathbb{N} \to \mathbb{P}$ given by a random sequence of natural numbers $\omega = \left(v_k(\omega) \mid k \in \mathbb{N}\right)$ where randomness of each term v_j is determined by the choice of elementary event $\omega \in \Omega$ due to a probability distribution P defined by a probability space $(\Omega, \mathcal{F}, P)$. Primes in

$\omega = (\nu_1, \nu_2 \ldots, \nu_j, \ldots)$ for each $\nu_k = k$ can be represented by the indicator function

$I_{\mathbb{P}}(n) = \xi(n)$ as a sequence of binary-valued variables $\xi(n) = \begin{cases} 1 \text{ if } n \in \mathbb{P} \\ 0, \text{ otherwise} \end{cases}$ for $n \in \mathbb{N}$

(see the Table 1.1 below).

This can be directly observed in the sequence of prime numbers below 100:

$$(2\ 3\ 5\ 7\ 11\ 13\ 17\ 19\ 23\ 29\ 31\ 37\ 41\ 43\ 47\ 53\ 59\ 61\ 67\ 71\ 73\ 79\ 83\ 89\ 97)$$

Table 1.1

The sequence $(\xi(n) \,|\, 1 \le n \le 100)$ *of sequential primes among natural numbers from 1 to 100 represented by values of n such that variables* $\xi(n) = 1$ if n is prime*:*

0 1 1 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 0 0 0

0 0 1 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0

The sequence consecutive primes $\nu_n = n \cdot \xi(n)$ *separated by zeros standing on the place of composite numbers*:

0 2 3 0 5 0 7 0 0 0 11 0 13 0 0 0 17 0 19 0 0 0 23 0 0

0 0 0 29 0 31 0 0 0 0 0 37 0 0 0 41 0 43 0 0 0 47 0 0 0

0 0 53 0 0 0 0 0 59 0 61 0 0 0 0 0 67 0 0 0 71 0 73 0

0 0 0 79 0 0 0 83 0 0 0 0 0 89 0 0 0 0 0 0 0 97 0 0 0

In Number Theory we are interested in recursive sequences of numbers, generated by certain recurrent relations, mostly nonlinear. Let $(\Omega, \mathcal{F}, P)$ be a probability space, where Ω is a set of all $\mathbb{N}$-valued sequences, $\mathcal{F}$ is a σ-algebra generated by the algebra of cylinder sets in Ω, and P is a probability measure on $(\Omega, \mathcal{F})$. From a probabilistic point of view , see [17] , a recursive sequence $\{\nu(n) \mid n \in \mathbb{N}\}$, in a more general setting, can be viewed as a realization of a *stochastic* (or *random*) *process* $\nu = \{\nu(t) \mid t \in T\}$ where T is interpreted as "discrete time" set (if $T \subseteq \mathbb{Z}$) or as "a continuous time" set (if $T \subseteq \mathbb{R}$). For every $t \in T$ random variables $\nu(t) = \nu(\omega, t)$ are $(X, \mathcal{B}) / (\Omega, \mathcal{F})$ - measurable functions $\nu(t) : \Omega \to X$, such that $(\nu(t))^{-1}(A) \in \mathcal{F}$ for any $A \in \mathcal{B}$. In other words, each mapping $(\nu(t))^{-1}$ is a homomorphisms of σ-algebra $\mathcal{B}$ into σ-algebra $\mathcal{F}$. The set T is considered as a one-parametric group ($\mathbb{Z}$ or $\mathbb{R}$) or as a one-parametric semigroup ($\mathbb{Z}^+$ or $\mathbb{R}^+$). The set X in a measurable space $(X, \mathcal{B})$ is called a *state space* of the process $\nu = \{\nu(t) \mid t \in T\}$, where

$$\nu \in X^T = \prod_{t \in T} X,\ \mathcal{B}^T = \bigotimes_{t \in T} \mathcal{B} .$$

In what follows we restrict our analysis to the case of the discrete time set

$$T = Z^+ = \mathbb{N} \cup \{0\}.$$

This is convenient to identify an elementary event $\omega \in \Omega$ with *a path* (or *trajectory*) $\omega(t) \equiv \nu(\omega, t) = x(t) \in X,\ t \in T$, and the probability space $(\Omega, \mathcal{F}, P)$ with its canonical realization $(X^T, \mathcal{B}^T, P_X)$ where a probability measure P_X is detetermined on all n-dimensional cylinder sets $C_{t_1, t_2, \ldots, t_n} = \left(x \in X^T \mid x_{t_1} \in A_1, x_{t_2} \in A_2, \ldots, x_{t_n} \in A_n\right)$ for

$A_k \in \mathcal{B}, k = 1, 2, \ldots, n;\ (t_1, t_2, \ldots, t_n) \in T^n, n \in \mathbb{N}$,

by the assignment of probabilities $P_X(C_{t_1, t_2, \ldots, t_n}) = P\left\{\bigcap_{k=1}^{n} \nu(t_k)^{-1}(A_k)\right\}$.

Following H. Furstenberg in [16] , we consider the *recurrence* as a central number-theoretic phenomenon to be studied. Then, we introduce a *measurable dynamical system* $(X, \mathcal{B}, \theta)$ as follows. Set $\theta = (\theta_t)_{t \in T}$ is a one-parametric *semigroup* (if $T \in \mathbb{N} \cup \{0\}$) or a *group* (if $T \in \mathbb{Z}$) of *transformations* θ_t acting on $(X, \mathcal{B})$

such that:

1. $\theta_t \nu = \nu(\omega, t)$ is a $\mathcal{F}/\mathcal{B}$-measurable function on the direct product $\Omega \times T$,
2. $\theta_0 \nu = \nu(\omega, 0)$ (if $0 \in T$),
3. $\theta_{t+s} = \theta_t \circ \theta_s$ (a semigroup property of transformation compositions).

In the framework of Probability Theory we consider basic sequences $\left(\xi(n) \mid n \in \mathbb{N}\right)$ as realizations of $(0,1)$-valued random variables traditionally called *Bernoulli variables*.

To avoid pure heuristic justification of probabilistic ,conclusions, we try to conduct our discourse entirely in the framework of Probability Theory. This means that, prior to discussion of dependence issues related to sequences like $\omega = \left(\nu_1, \nu_2 \ldots, \nu_j, \ldots\right)$, we should introduce random variables $\nu_j : \Omega \to \mathbb{N}$ with the corresponding probability distribution P_j defined on σ-algebra $\mathcal{F}_j$ of events $\nu_j^{-1}(A)$ (generated in our context by all subsets $A \subseteq \mathbb{N}$).

We assume that a binary-valued sequence $\left(\xi(n) \mid n \in \mathbb{N}\right)$, where $\xi(n) = \begin{cases} 1, \text{ if } n \in \mathbb{P} \\ 0, \text{ otherwise} \end{cases}$, representing primes is a realization of a non-stationary sequence of possibly *dependent* Bernoulli variables, by postulating probabilities

$$P\left\{\xi(n) = 1\right\} = p_n, P\left\{\xi(n) = 0\right\} = 1 - p_n \text{ where } 0 < p_n < 1. \tag{1.3}$$

The major challenges in the study of such sequences are evaluation of q_n in (2) and analysis of dependence of random variables $\left(\xi(n) \mid n = 1, 2, \ldots\right)$ included in the sequence. The problem of dependence of events and random variables in the framework of Number Theory had been discussed in some detail in the monograph of Mark Kac [4]. In number of works authors tried to avoid a rigorous probabilistic approach based on the concept of sigma-additive probability measures and the corresponding probability spaces, and considered instead so-called 'density' measures, which are additive but not sigma-additive. As M. Kac underlined in [4], the concept of independence "though of central importance in probability theory, is not a purely mathematical notion", and it appears quite naturally in Statistical Physics. He mentioned that "the rule of multiplication of probabilities of independent events is an attempt to formalize this notion and to build a calculus around it". Moreover, the notions of statistical (probabilistic) independence and dependence of events have been sometimes confused with the mathematical (functional) or logical dependence. Both dependence and independence of "events" in Number Theory are results of complicated recursive nonlinear relations between terms of numeric sequences, which can generate a 'dynamical chaos', imitating pseudo-randomness in the long run behavior of such (theoretically) purely 'deterministic' sequences. The precise prediction of behavior of terms in the sequences demands almost impossible calculations based on expanding memory of prehistory of their evolution. To make a study feasible and overcome 'the curse of dependence', a typically suggested heuristic assumption is that terms in $\left(\xi(n)\right)_{n\in\mathbb{N}}$ are asymptotically independent, or uncorrelated, or 'weakly' dependent in a certain sense. In the framework of modified Cramér's model we show that the sequence of dependent random variables $\left(\xi(n)\right)_{n\in\mathbb{N}}$ is *asymptotically pairwise Bernoullian* (that is asyptotically pairwise independent) in a sense that we are going to discuss below.

Surprisingly, in many discussions of probabilistic interpretations of Number Theory problems, some authors use 'by default' an approach as in the following sentence [1]: "Assume that we choose number X at random from 1 to n.

Then $\text{Prob}(X \text{ is prime}) = \dfrac{\pi(n)}{n}$ …"

The above sentence, due to its ambiguity, raises the following comments and objections.

1) When one chooses number X "at random", it is presumed that the probability distribution of X is known (at least theoretically). The formula $\text{Prob}(X \text{ is prime}) = \dfrac{\pi(n)}{n}$ tells us that the probability distribution is assumed to be uniform on the sequence of integers $\{1,2,3,\ldots,n\}$. Here $\pi(n) = \#\{p \in \mathbb{P} \mid p \le n\}$ is a counting function of number of primes not exceeding n. If the probability distribution of X is not uniform on the interval of integers $[1,n] = \{1,2,\ldots,n\}$, then, in a statistical framework, $\frac{\pi(n)}{n}$ can be interpreted not as a probability but rather as a relative frequency of occurrences of prime numbers in the interval $[1,n]$. One of goals in our study is to construct a probabilistic model for the "statistical" distribution of primes given by the observed frequencies $\frac{\pi(n)}{n}$. Notice again the obvious fact that a *discrete uniform probability distribution* does not exist on an infinite support, that is on infinite subsets of $\mathbb{N}$ (including $\mathbb{N}$ itself).
2) The following analysis is about divisibility of ν by a prime $p \le n$. Denote $p \cdot \mathbb{N}$ a set of all multiples of number p. As mentioned above, the probability $P\{\nu \in p \cdot \mathbb{N}\}$ does not exists if ν is evenly distributed on $\mathbb{N}$. But the problem can be easily resolved if one refers the probability $P\{\nu \in p \cdot \mathbb{N}\}$ to the class $C_{p,0} = \{n \mid n = k \cdot p,\ k \in \mathbb{N}\}$ of integers in $\mathbb{N}$ congruent 0 modulo p. There are exactly p congruent classes modulo p: $C_{p,r} = \{n \mid n = k \cdot p + r; 0 \le r \le p-1; k \in \mathbb{N} \cup \{0\}\}$, which make a partition of $\mathbb{N}$.

Then, we can define a probability distribution $P\left(C_{p,r}\right)=q_{p,r}$ $(r=0,1,2,\ldots,p-1)$ on $\left\{C_{p,0},C_{p,1},\ldots,C_{p,p-1}\right\}$ such that $\sum_{r=0}^{p-1}q_{p,r}=1$. Equal probabilities to randomly choose a class of congruence for a number ν given by $P\left(C_{p,r}\right)=q_p$ for all r: $0\le r\le p-1$, imply $P\left(C_{p,r}\right)=\frac{1}{p}$. Particularly, we have $C_{p,0}=p\cdot\mathbb{N}$, so that

$$P\left\{\nu\in p\cdot\mathbb{N}\right\}=P\left(C_{p,0}\right)=\frac{1}{p} \quad . \qquad (2.1.22)$$

Considering $P\left\{\nu\in p\cdot\mathbb{N}\right\}$, we assumed that random variable ν can take any value within $p\cdot\mathbb{N}$. The value of the probability can be different from $P\left\{\nu\in p\cdot\mathbb{N}\right\}=\frac{1}{p}$ if we impose some limitations on ν, say, if we assume that $\nu\le n$. For arbitrary $n\in\mathbb{N}$ and a given probability distribution of ν, an event $\left\{\nu\le n\right\}$ may not belong, in general, to the algebra of events created by the partition of $\mathbb{N}$ into p congruence classes $\left\{C_{p,r}\mid r=0,1,2,\ldots,p-1\right\}$, and it would be impossible to assign a probability value to the event $\left\{X\in C_{p,0}\cap[1,n]\right\}$, where we denote $[1,n]=\left\{k\mid k=1,2,\ldots,n\right\}$.

Since $[1,n]$ is a finite set, we can define a uniform probability distribution on this set, but the agreement of this distribution with the assumption $P\left\{C_{p,0}\cap[1,n]\right\}=\frac{1}{p}$ would depend on the choice of n, specifically, on divisibility of n by p.

For example, if $p=3$ and $n=20$, we have $C_{3,0}\cap[1,20]=\{3,6,9,12,15,18\}$ and $P\left\{C_{3,0}\cap[1,20]\right\}=\frac{6}{20}=0.3\ne\frac{1}{3}$.

For $p=3$ and $n=21$ we have $C_{3,0}\cap[1,21]=\{3,6,9,12,15,18,21\}$ so that

$P\{C_{3,0}\cap[1,21]\}=\frac{7}{21}=\frac{1}{3}=0.333\ldots$.

3) Independence of divisibility of random number ν by different primes is determined by the choice of probability distribution of ν. As it had been noticed by Mark Kac in [4], "primes play a game of chance". He pointed out to the obvious fact that ν to be divisible by both primes p and q is equivalent of being divisible by $p\cdot q$. This mean that if $P\{C_{m,0}\}=\frac{1}{m}$ for any positive integer m, then, since $C_{p\cdot q,0}=C_{p,0}\cap C_{q,0}$, we have

$$P\{C_{pq,0}\}=P\{C_{p,0}\}\cdot P\{C_{q,0}\} \text{ because } \frac{1}{p\cdot q}=\frac{1}{p}\cdot\frac{1}{q}. \qquad (2.1.23)$$

Mark Kac was not able to establish and use the independence of divisibility events in terms of a probability theory since he used a density set functions $d(A)=\lim_{n\to\infty}\frac{A(n)}{n}$, where $A(n)=A\cap[0,n]$, $A\subset\mathbb{N}$, which is additive but not σ-additive.

Definition. 1.1

We call a probability distribution P_f on $\mathbb{N}$ *multiplicative or completely multiplicative* if for all $A\subseteq\mathbb{N}$ we have:

$$P_f\{X\in A\}=\frac{1}{Z}\sum_{n\in A}f(n)\text{ , where } f:\mathbb{N}\to(0,1], \qquad (2.1.24)$$

is a *multiplicative*, *or respectively, completely multiplicative function,*

such that $Z=\sum_{n\in\mathbb{N}}f(n)$ is a convergent series.

As we show below, independence of divisibility of random number X by different primes can be guaranteed if X has a multiplicative probability distribution defined above.

Each prime number p determines a partition of the set $\mathbb{N}$ into p classes of congruence modulo p: $C_{p,r}$, where $r \in \{0,1,2,\ldots,p-1\}$. We show below that a randomly chosen value ν with the multiplicative distribution P_f is divisible by natural m with probability $f(m)$. For $f(n)=\dfrac{1}{n^s}$ $(s>1)$ the probability P_f on $\mathbb{N}$ is Zeta probability distribution, and random ν with Zeta distribution is divisible by a prime number p with probability $\dfrac{1}{p^s}$ so that for each $p \in \mathbb{P}$,

$$P_s\left\{\nu \in C_{p,0}\right\} = \frac{1}{p^s} \ , P_s\left\{\nu \notin C_{p,0}\right\} = 1-\frac{1}{p^s} \tag{2.1.25}$$

Each natural n, due to the Fundamental Theorem of Arithmetic, can be represented in the unique form

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} = \prod_{j=1}^{k} p_j^{\alpha_j} \tag{2.1.26}$$

where $p_1, p_2, \ldots, p_k$ are distinct primes, and $a_1, a_2, \ldots, a_k$ are natural numbers. Formula (2.1.26) is called a *canonical representation* of n where $a_1, a_2, \ldots, a_k$ are called *multiplicities* of prime factors of n.

Thus, the probability that p does not divide ν equals $P_s\left\{\alpha(\nu,p)=0\right\} = 1-\dfrac{1}{p^s}$.

In general, the event $\left\{\alpha(\nu,p)=k\right\}$ in (5.1) means that p^k divides ν but p^{k+1} does not divide ν :

$$P_s\left\{\alpha(\nu,p)=k\right\} = \left(\frac{1}{p^s}\right)^k \cdot \left(1-\frac{1}{p^s}\right), \ k=0,1,2,3,\ldots \tag{2.1.26}$$

and we have

$$E\alpha(\nu,p)=\frac{p^{-s}}{1-p^{-s}}=\frac{1}{p^{s}-1};\ Var\big(\alpha(\nu,p)\big)=\frac{p^{-s}}{\left(1-p^{-s}\right)^{2}}=\left(\frac{p^{s}}{p^{s}-1}\right)\cdot\left(\frac{1}{p^{s}-1}\right).$$

Sum $\varphi(\nu)=\sum_{p\in\mathbb{P}}\alpha(\nu,p)$ counts the total number of prime factors (with multiplicities) in the prime factorization of $\nu=\prod_{p\in\mathbb{P}}p^{\alpha(\nu,p)}$. Here are parameters of $\varphi(\nu)$:

$$E_{s}\left[\varphi(\nu)\right]=\sum_{p\in\mathbb{R}}\frac{1}{p^{s}-1},\ Var_{s}\left[\varphi(\nu)\right]=\sum_{p\in\mathbb{R}}\left(\frac{p^{s}}{p^{s}-1}\right)\cdot\left(\frac{1}{p^{s}-1}\right).$$

Assume now that there is a vector $\vec{p}=\left(p_1,p_2,\ldots,p_N\right)$, which components are N different prime numbers, and we consider a multiplicative semigroup $S(\vec{p})$ with unity, generated by components of vector $\vec{p}$ and $\{1\}$.

For any $n\in S(\vec{p})$ we have $n=\prod_{j=1}^{k}p_j^{\alpha_j}$ where $\alpha_j>0$ for all j $(1\le j\le k),\ k\le N$.

By using computer simulation, we can generate N pseudo-random variables $\alpha_j=\alpha(p_j,\nu),\ 1\le j\le N$, where each $\alpha(p_j,\nu)$ has a geometric distribution with parameter $\dfrac{1}{p_j}$ and then, simulate a ‘pseudo-random’ number $\nu=\prod_{j=1}^{k}p_j^{\alpha_j}$ with $k=k(\nu)\ \le\ N$.

Further we consider a multiplicative semigroup $S(\mathbb{P}_N)$ generated by all primes not exceeding N, that is $\mathbb{P}_N=\mathbb{P}\cap\left\{p\le N\mid p\in\mathbb{P}\right\}$.

THEOREM 1.1.

If P_f is a *multiplicative probability distribution* on $\mathbb{N}$ and ν is a random variable such that

$$P_f\{\nu \in A\} = \frac{1}{Z}\cdot\sum_{n\in A} f(n) \text{ where } A\subseteq\mathbb{N},\ f:\mathbb{N}\to(0,1]\ ,$$

then

1) For any natural $m\geq 2$ random event E of occurrence of a random number ν divisible by m has probability $P_f(E)=P_f(C_{m,0})=f(m)$.

2) for any two mutually prime numbers m_1 and m_2, random events E_1 and E_2 of occurrence of ν divisible by both m_1 and by m_2, respectively, are P_f-independent events: $P_f(E_1\cap E_2)=P_f(E_1)\cdot P_f(E_2)$. Since $E_1=C_{m_1,0}$, $E_2=C_{m_2,0}$ and $E_1\cap E_2=C_{m_1\cdot m_2,0}$ we have, equivalently,

$$P_f(C_{m_1,0}\cap C_{m_2,0})=P_f(C_{m_1,0})\cdot P_f(C_{m_2,0})$$

Proof.

For $m=m_1\cdot m_2$ we have:

$$P_f\left(C_{m,0}\right)=\frac{1}{Z}\sum_{k\in\mathbb{N}} f(m\cdot k)=\frac{1}{Z}\sum_{k\in\mathbb{N}} f(m)\cdot f(k)=f(m)=f(m_1)\cdot f(m_2) \quad \text{since } \frac{1}{Z}\sum_{k\in\mathbb{N}} f(k)=1,$$

and $P_f\left(C_{m_i,0}\right)=\frac{1}{Z}\sum_{k\in\mathbb{N}} f(m_i\cdot k)=f(m_i)\ (i=1,2)$. Then, $C_{m_1\cdot m_2}=C_{m_1}\cap C_{m_2}$ implies

$$P_f(C_{m_1,0}\cap C_{m_2,0})=P_f(C_{m_1\cdot m_2,0})=P_f(C_{m_1,0})\cdot P_f(C_{m_2,0})$$

Q.E.D.

The following theorem states that the assumption that the probability distribution P_f on $\mathbb{N}$ is 'complete multiplicative' (with an appropriate choice of function f) is necessary and sufficient condition for such distribution P_f to be Zeta distribution.

THEOREM 1.2.

Let ν be a random variable with values in $\mathbb{N}$ with probability distribution

$$P_f\{\nu \in A\} = \frac{1}{Z}\sum_{n\in A} f(n) \ , \qquad (2.1.28)$$

where $f:\mathbb{N}\to[0,1]$, $A\subseteq\mathbb{N}$ and $Z=\sum_{n=1}^{\infty} f(n)$ is a convergent series.

The series $Z=\sum_{n=1}^{\infty} f(n)$ takes a form of the 'Euler product of the series' [12, p.230]:

1) if f in (2.1.28) is multiplicative, then

$$Z=\sum_{n=1}^{\infty} f(n)=\prod_{p\in\mathbb{P}}\left[1+f(p)+f(p^2)+\cdots\right]$$

2) if f in (2.1.28) is a completely multiplicative function such that

$0<f(p)<1$ for all $p\in\mathbb{P}$, then

$$Z=\sum_{n=1}^{\infty} f(n)=\prod_{p\in P}\frac{1}{1-f(p)}$$

3) the probability distribution P_f is a Riemann Zeta distribution

$P_{\zeta(s)}\{\nu=n\}=\dfrac{n^{-s}}{\zeta(s)}$, $n\in\mathbb{N}$, for any choice of $s>1$. Further we denote $P_{\zeta(s)}=P_s$.

Proof.

1) Let $S(\mathbb{P}_N)$ be a semigroup of all integers generated by $\mathbb{P}_N\cup\{1\}$,

$\mathbb{P}_N=\{p \mid p\le N, p\in\mathbb{P}\}$.

Due to the Fundamental Theorem of Arithmetic,

$$n=\prod_{p\in\mathbb{P}^*} p^{\alpha(n,p)},\ \text{where } \alpha(n,p)\ge 0,\ \alpha(n,p)=\begin{cases} a_j>0 \text{ if } p^{a_j}\mid n \text{ and } p^{a_j+1}\nmid n\\ 0, \text{otherwise}\end{cases}$$

Then, if f is a multiplicative function, we have

$$Z=\sum_{n=1}^{\infty} f(n)=\sum_{n=1}^{\infty}\left[\prod_{p\in\mathbb{P}} f(p^{\alpha(n,p)})\right]=\prod_{p\in\mathbb{P}}\left[\sum_{k=0}^{\infty} f(p^k)\right]=\prod_{p\in\mathbb{P}}\left[1+f(p)+f(p^2)+\cdots\right]$$

2) In the proof above we have used the multiplicative property of function f.

If f is completely multiplicative, we have $f(p^k)=\left(f(p)\right)^k$. Then, we can write

$1+f(p)+\left(f(p)\right)^2+\left(f(p)\right)^3+\cdots=\dfrac{1}{1-f(p)}$ and the above equality takes a form:

$$Z=\sum_{n=1}^{\infty}f(n)==\prod_{p\in\mathbb{P}}\left[\sum_{k=0}^{\infty}\left(f(p)\right)^k\right]=\prod_{p\in\mathbb{P}}\frac{1}{1-f(p)}$$

Notice that the right-hand sides of the above equalities are convergent infinite products, since the left-hand side is given by the convergent series.

3) Notice that for any $n\in S(\mathbb{P}_N)$ we have $n^s=\prod_{p\in\mathbb{P}}p^{\alpha(p)\cdot s}$, where $\alpha(p)\in\mathbb{N}\cup\{0\}$.

Since $\dfrac{1}{1-\dfrac{1}{p^s}}=\sum_{k=0}^{\infty}\left(\dfrac{1}{p^s}\right)^k$, we have $\zeta_{\mathbb{P}_N}(s)=\prod_{p\le N}\left[\sum_{k=0}^{\infty}p^{-sk}\right]=\sum_{\{a(p),p\le N\}}\prod_{p\le N}p^{-s\cdot a(p)}=\sum_{n\in S(\mathbb{P}_N)}n^{-s}$

Denote $\xi_p(\nu)=p^{\alpha(\nu,p)}$, $\xi=p$ .Then, $P\left\{\xi_{\nu,p}=p^{\alpha(\nu,p)}\right\}=\left[P\left\{\xi=p\right\}\right]^{\alpha(\nu,p)}$

For any natural m we denote the event "m divides ν" as $E=\{m|\nu\}$ and the opposite event "m does not divide ν" as $\bar{E}=\{m\nmid\nu\}$. The probability that a prime number p divides ν is $P\{p|\nu\}=f(p)$ and the probability that p does not divide ν is $P\{p\nmid\nu\}=1-f(p)$. The probability that the number ν divides p^k and does not divide p^{k+1} is given by the formula

$$P\left\{\left(p^k|\nu\right)\cap\left(p^{k+1}\nmid\nu\right)\right\}=\left(f(p)\right)^k\cdot\left(1-f(p)\right)$$

Then, by virtue of Theorem 2.3 and the canonical factorization of n, we have

$$\begin{aligned}P\{\nu=n\}&=\prod_{p\in\mathbb{P}}P\left\{\left(p^{\alpha(n,p)}|\nu\right)\cap\left(p^{\alpha(n,p)+1}\nmid\nu\right)\right\}\\&=\prod_{p\in\mathbb{P}}\left[\left(f(p^{\alpha(n,p)})\right)\cdot\left(1-f(p)\right)\right]=\prod_{p\in\mathbb{P}}\left[f(p)\right]^{\alpha(\nu,p)}\cdot\prod_{p\in\mathbb{P}}\left[1-f(p)\right]\end{aligned}\qquad(2.1.29)$$

Summation of both sides of (2.29) results in the formula:

$1=\sum_{n\in} P\{\nu=n\}=\prod_{p\in P}(1-f(p))\cdot\sum_{\nu\in}\prod_{p\in P}[f(p)]^{\alpha(\nu,p)}$, which implies:

$$\prod_{\nu\in\mathbb{N}}\frac{1}{1-f(p)}=\sum_{\nu\in\mathbb{N}}\prod_{p\in\mathbb{P}}f(p^{\alpha(n,p)})=\sum_{\nu\in\mathbb{N}}f\left(\prod_{p\in\mathbb{P}}p^{\alpha(n,p)}\right)=\sum_{\nu\in\mathbb{N}}f(n)=Z\ , \qquad (2.1.30)$$

provided that $f(n)$ is such that the infinite product and the infinite sum in the above formulas are both convergent. Completely multiplicative function $f:\mathbb{N}\to(0,1]$ satisfies the functional equation $f(x\cdot y)=f(x)\cdot f(y)$, known as one of 'fundamental' Cauchy functional equations and due to Theorem 3, p.41 in [13], for positive x,y has the most general solution of the form $f(x)=e^{c\cdot\ln x}=x^c$. Obviously, in our context $f(n)=n^{-s}(s>1)$ is a completly multiplicative arithmetic function and for this choice of f, $Z(f)=\zeta(s)$ is Zeta function which generates Zeta distribution

$$P_{\zeta(s)}\{\nu=n\}=\frac{1}{n^s\cdot\zeta(s)},\ n\in\mathbb{N}.$$

Q.E.D.

Remark 1.1.

The problem with the choice $f(n)=\dfrac{1}{n}$ for $s=1$ is that it leads to the divergent harmonic series $\zeta(1)=\sum_{n=1}^{\infty}\dfrac{1}{n}$. To avoid the situation with the series divergence, we follow the steps of Euler [3] by restricting values of s to $s>1$. Zeta function $\zeta(s)=\sum_{n=1}^{\infty}\dfrac{1}{n^s}$ is well known to be directly related to the probability distribution of prime numbers.

This motivates the choice of Zeta distribution. If p divides ν, then $\nu=p\cdot\nu'$ while the quotient $\nu'=\dfrac{\nu}{p}$ is again distributed over p classes of congruence $C_{p,r}$, and so on. An odd number $\nu \le n$ is prime if and only if it does not divide all primes less than or equal to $\sqrt{n}$:

$$P_s\{\nu \in \mathbb{P} | \nu \le \mathrm{n}\} = P_s\left\{\bigcap_{p \le \sqrt{n}} [\alpha(\nu, p) = 0]\right\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right) \qquad (2.1.31)$$

Since for $p \in \mathbb{P}$ we have $P_s\{\alpha(\nu, p) = k\} = p^{-sk} \cdot \left(1 - p^{-s}\right)$ for all $k = 0, 1, 2, \dots$.

In particular, $P_s\{\alpha(\nu, p) = 1\} = p^{-s} \cdot \left(1 - p^{-s}\right)$, and $P_s\{\alpha(\nu, p) = 0\} = 1 - p^{-s}$.

Then, the probability of. $\{\nu = p_j \in \mathbb{P}\}$ is calculated as

$$P\{\nu = p_j\} = P_s\{\alpha(\nu, p_1) = 0, \dots, \alpha(\nu, p_{j-1}) = 0, \alpha(\nu, p_j) = 1, \alpha(\nu, p_{j+1}) = 0, \dots\}$$

$$= \left(\frac{1}{p_j^s}\right) \cdot \left(1 - \frac{1}{p_j^s}\right) \cdot \prod_{k \ne j} P_s\{\alpha(\nu, p_k) = 0\} = \left(\frac{1}{p_j^s}\right) \cdot \prod_{k=1}^{\infty} \left(1 - \frac{1}{p_j^s}\right) = \frac{p_j^{-s}}{\zeta(s)}$$

In general, for any natural number $\nu = n = \prod_{p \in \mathbb{P}} p^{\xi_p} = p_1^{k_1} \cdot p_2^{k_2} \cdots p_m^{k_m} \cdots$, we have

$$P_s\{\nu = n\} = \prod_{p \in \mathrm{P}} \left[\left(\frac{1}{p^s}\right)^{\alpha(n,p)} \cdot \left(1 - \frac{1}{p^s}\right)\right] = \prod_{p \in \mathrm{P}} \left(\frac{1}{p^s}\right)^{\alpha(n,p)} \cdot \prod_{p \in \mathrm{P}} \left(1 - \frac{1}{p^s}\right) = \frac{1}{n^s} \cdot \zeta^{-1}(s),$$

that is

$$P_s\{\nu = n\} = \frac{n^{-s}}{\zeta(s)} \qquad (2.1.32)$$

Formula (2.1.32) may provide some probabilistic interpretations of Riemann Zeta function.

If ν has Zeta probability distribution, then the probability that $\nu(\omega)$ for certain ω results in a prime number is evaluated as

$$P_s\{\nu \in \mathbb{P}\} = \frac{1}{\zeta(s)} \sum_{p \in \mathbb{P}} p^{-s} \qquad (2.1.33)$$

Taking into account the recursive dependence of the sequence of prime numbers $p \le \nu$ with the memory size $\sqrt{\nu}$, we have the identity of the events $\{\nu \in \mathbb{P}\} = \bigcap_{p \le \sqrt{\nu}} \{p \nmid \nu\}$, which means that an *odd number* $\nu = n$ is prime if and only if any prime $p \le \sqrt{n}$ does not divide n. Formally:

$$r = \mathrm{mod}(n, p') \ne 0 \text{ for all primes } p' \le \sqrt{n} .$$

This implies:

$$P_s\{\textit{odd } \nu \text{ is prime }\} = \prod_{p \le \sqrt{\nu}} \left(1 - \frac{1}{p^s}\right) \qquad (2.1.34)$$

Since $\{\nu \le n\} = \bigcup_{i=1}^{n} \{\nu = i\}$, we have

$$P_s\{\nu \le n\} = \frac{\sum_{k=1}^{n} k^{-s}}{\zeta(s)} \text{ and } P_s\{\nu \in \mathbb{P} \text{ and } \nu \le n\} = \frac{\sum_{p \le n} p^{-s}}{\zeta(s)} .$$

We could compare the last probability with the frequency estimate $\frac{\pi(n)}{n}$ or with the Cramér's model prediction $\frac{1}{\ln n}$, though, dependence of probability P_s on parameter $s > 1$ makes the above formulas much harder to interpret. As we know, one can circumvent divergence of $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ for $s \le 1$ by using the analytic continuation of $\zeta(z)$ on the complex plane $\mathbb{C}$, as suggested by Riemann. Meanwhile, as we have mentioned above, the use of *Incomplete Product Zeta* function (*IPZ*) $\zeta_{\mathbb{P}_N}(s)$ defined as a partial product of $\zeta(s)$,

$\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N} \frac{1}{1 - \frac{1}{p^s}}$, provides another opportunity to deal with the divergence of $\zeta(1)=\prod_{p\in\mathbb{P}}\frac{1}{1-\frac{1}{p^s}}$ for $s=1$.

LEMMA 1.2.

Let $S(\mathbb{P}_N)$be a semigroup of all integers generated by $\mathbb{P}_N\cup\{1\}$,

$\mathbb{P}_N=\{p\mid p\le N, p\in\mathbb{P}\}$.

Then,

$$\zeta_{\mathbb{P}_N}(s)=\sum_{n\in S(\mathbb{P}_N)} n^{-s}.$$

Proof.

Notice that for any $n\in S(\mathbb{P}_N)$ we have $n^s=\prod_{p\in\mathbb{P}} p^{\alpha(p)\cdot s}$, where $\alpha(p)\in\mathbb{N}\cup\{0\}$.

Since $\frac{1}{1-\frac{1}{p^s}}=\sum_{k=0}^{\infty}\left(\frac{1}{p^s}\right)^k$, we have $\zeta_{\mathbb{P}_N}(s)=\prod_{p\le N}\left[\sum_{k=0}^{\infty}p^{-sk}\right]=\sum_{\{a(p),p\le N\}}\prod_{p\le N}p^{-s\cdot a(p)}=\sum_{n\in S(\mathbb{P}_N)}n^{-s}$.

Q.E.D.

We consider now the corresponding probability distribution $P_{s,N}$, $s>1$, and find the probability of an odd ν to be a prime number in the set of numbers $S(\mathbb{P}_N)$ generated by primes not exceeding N:

$$P_{s,N}\{\text{odd } \nu \text{ is prime} \mid \nu\in \mathrm{S}(\mathbb{P}_N)\}=\prod_{p\le\sqrt{\nu}}\left(1-\frac{1}{p^s}\right) \tag{2.1.35}$$

2. Multiplicative and Additive Recurrent models for Primes

The famous Harald Cramér's model [2,3] describes the occurrence of prime numbers as a *sequence of independent Bernoulli variables* with probabilities

$$P\{\xi(n)=1\}=\frac{1}{\ln n},\ \ P\{\xi(n)=0\}=1-\frac{1}{\ln n},\ \text{ where } n\geq 2. \tag{2.1}$$

In what follows, we provide rigorous arguments in support of some aspects of Cramér's model, especially related to the values of probabilities $q_n=\frac{1}{\ln n}$, and then analyze dependence of $\xi(n)$ in the sequence $(\xi(n)\,|\,n=1,2,\ldots)$. As we have discussed above, appearance of a prime $n=\nu(k)$ in the sequence $\{\nu(k)=k\,|\,k\in\mathbb{N}\}$ are dependent events determined by the prehistory $\mathcal{F}_{\sqrt{n}}=\sigma\{\nu(k)\,|\,1\leq k\leq\sqrt{n}\}$.

Obviously, if $\nu(k)=p\in\mathbb{P}$, then $\nu(k+1)=p+1\notin\mathbb{P}$ since $p+1$ is an even number. Even if we restrict values of $\nu(k)$ by odd numbers $2k+1$, still divisibility of $\nu(2k+1)=n$ by the previously occurred primes would depend on the prehistory $\mathcal{F}_{\sqrt{n}}$. Therefore, the sequence of consecutive primes and the corresponding Bernoulli variables $\xi(n)$ cannot be interpreted as occurrence of independent events in the sequence, or as a realization of a Markov chain with a constant size of 'memory', because for each $\nu(k)=n$ the size $[\sqrt{n}]$ of the 'memory' $\mathcal{F}_{\sqrt{n}}$ increases in the sequence with n. We analyze the sequence of prime numbers $\{\nu(k)=p\,|\,p\in\mathbb{P},k\in\mathbb{N}\}$ by using *multiplicative and additive models*.

In any kind of a model, we will be using the equivalent canonical realizations

$$(\Omega,\mathcal{F},P)=\left(X^T,\mathcal{B}^T,P_X\right)\ \text{ so that }\ \nu(\omega,t)=\nu(t)\ .$$

The transformations $\theta_t: X^T\to X^T, t\in T$, are $\mathcal{B}^T/\mathcal{B}^T$ -measurable.

We define the transformations by $\theta_s\nu(t)=\nu(t+s)$, for $s,t\in T$.

A *multiplicative model* is based on the *canonical representation of primes* [5, p.18]:

$$n = \prod_{p \in \mathbb{P}} p^{\alpha(n,p)} \text{ where } \alpha(n,p) = \begin{cases} \alpha_p > 0 \text{ if } p \text{ divides } n \\ 0, \text{ otherwise} \end{cases} \tag{2.2}$$

and is concerned with the questions of divisibility of integer-valued random variables by integers, and with their connection with the Zeta probability distribution:

$$P_s\{v \in A\} = \frac{1}{\zeta(n)} \cdot \sum_{n \in A} \frac{1}{n^s} \text{ , for any subset } A \subseteq \mathbb{N}. \tag{2.6}$$

For the multiplicative model of the dynamical system representing (5), where $v = n$, we define

$$\theta_i v = v(i); \; v(0) = 1, \theta_{i+1} v = \theta_i v \cdot \eta(i+1),$$

$$\text{where } \eta(n+1) = p_{n+1}^{\alpha_{n+1}(v)} (i = 0,1,2,\ldots,\kappa(v)-1).$$

See (9) in more detailed discussion.

Additive models are useful in problems related to counting of various types of integers in $\mathbb{N}$. In additive models dynamical systems are defined by the equations:

$$\theta_i v = v(i); v(0) = 0, \theta_{i+1} v = \theta_i v + \xi(i+1),$$

where definition of the 'updating' term $\xi(n+1)$ determines the specifics of the model, as illustrated below. First, we consider the function $\pi(x)$, counting the number of primes less than or equal to x. For this case, the updating term $\xi(n+1)$ is defined by the formula (1.3).

Second, we denote $\pi_d(x)$ the number of d-primes. A prime p is called a *d-prime* if the gap between p and its consecutive prime p' equals d, that is $p' - p = d$ and there are no primes between p and p'.

Third, for all $m \geq 3$ we consider the number $G(2m)$ of *Goldbach m-primes,* or *G_m-primes*, which are such primes p that a difference $2m - p$ is again a prime number.

In the first situation we use recurrent equations:

$$\begin{cases} \pi(1) = 0 \\ \pi(n+1) = \pi(n) + \xi(n+1), \ n \in \mathbb{N} \end{cases} \tag{2.7}$$

It is well-known that the connections between additive and multiplicative properties of numbers are extraordinarily complicated, and this leads to various difficult problems in Number Theory. We start from the *division algorithm* [5, p.19]. Given integer n and $m > 0$ there exists a unique pair of integers k and r such that $n = mk + r$, with $0 \le r \le m$. In this equation, $r = 0$ if and only if m divides n. We derive here a recursive formula generating a sequence of prime numbers: $2, 3, 5, 7, \ldots$ For any prime number $p \in \mathbb{P}$ and a natural number $n \ge 2$, consider a function $\mathrm{mod}(n, p) = r$ of residuals (remainders) such that $n = m \cdot p + r,\ 0 \le r < p,$ where $m \in \mathbb{N} \cup \{0\}$. Consider a vector of consecutive prime numbers $\vec{p}(n) = (p_1, p_2, \ldots, p_k)$ such that $p_k \le n$ and $p_{k+1} > n$. Index k determines here the value $\pi(n) = k$ for the number of primes less than or equal to n so that $\vec{p}(n) = (p_1, p_2, \ldots, p_{\pi(n)})$. For each coordinate p_i of vector the $\vec{p}(n)$ we determine the residual value $r_i = \mathrm{mod}(n, p_i),\ i = 1, 2, \ldots, \pi(n),$ and consider the vector of residuals $\vec{r}(n) = (r_1, r_2, \ldots, r_{\pi(n)})$. Notice that, due to the Sieve Algorithm, for an integer $n > 2$ to be prime it is necessary and sufficient that the all coordinates r_i of the 'reduced' vector of residuals $\vec{r}(n)$ such that $1 \le i \le \pi(\sqrt{n})$ be different from zero. Thus, the events $\left\{ \min\limits_{i \le \pi(\sqrt{\nu})} \{ r_i \mid r_i = \mathrm{mod}(\nu, p_i) \} > 0 \right\}$ and $\{\nu \in \mathbb{P}\}$are equivalent. See calculations below in the Table 2.1.

Table 2.1. *The recursive sequence of primes driven by their residuals*

n	$\pi(n)$	$\vec{p}(n)=\left(p_1,p_2,\ldots,p_{\pi(n)}\right)$	$\vec{r}(n)=\operatorname{mod}\left(n,\vec{p}(n)\right)=\left(r_1,r_2,\ldots,r_{\pi(n)}\right)$
2	1	(2)	(0)
3	2	$(2,3)$	$(1,0)$
4	2	$(2,3)$	$(0,1)$
5	3	$(2,3,5)$	$(1,2,0)$
6	3	$(2,3,5)$	$(0,0,1)$
7	4	$(2,3,5,7)$	$(1,1,2,0)$
8	4	$(2,3,5,7)$	$(0,2,3,1)$
9	4	$(2,3,5,7)$	$(1,0,4,2)$
10	4	$(2,3,5,7)$	$(0,1,0,3)$
11	5	$(2,3,5,7,11)$	$(1,2,1,4,0)$
12	5	$(2,3,5,7,11)$	$(0,0,2,5,1)$
13	6	$(2,3,5,7,11,13)$	$(1,1,3,6,2,0)$
…	…	…	…
30	10	$(2,3,5,7,11,13,17,19,23,29)$	$(0,0,0,2,8,4,13,11,7,1)$
31	11	$(2,3,5,7,11,13,17,19,23,29,31)$	$(1,1,1,3,9,5,14,12,8,2,0)$

We evaluate $P\left\{v\in\mathbb{P}\,|\,v=n\right\}$ assuming that a random integer v follows Zeta probability distribution. To assign a probability value to a set $m\cdot\mathbb{N}$ ("all multiples of number m"), we should refer it to the class $C_{m,0}=\left\{n\,|\,n=k\cdot m,\,k\in\mathbb{N}\right\}$ of integers in $\mathbb{N}$ congruent 0 modulo m so that $C_{m,0}=m\cdot\mathbb{N}$. There are exactly m congruent classes modulo m: $C_{m,r}=\{n\,|\,n=r+k\cdot n,\,k\in\mathbb{N}\cup\{0\}\},\,0\le r\le m-1$, which make a finite partition of $\mathbb{N}$. Then, for each integer $m>1$ we can define a probability distribution on $\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\}$:

$$P\{\nu\in C_{m,r}\}=q_{m,r}\ge 0,\,0\le r\le m-1 \quad \text{and} \quad \sum_{r=0}^{m-1}q_{m,r}=1,\,m=2,3,4\ldots$$

Theorem 2.1

Let ν be a random variable with Zeta probability distribution P_s and

$$\nu=\prod_{p\in\mathbb{P}}p^{\alpha(\nu,p)}=\prod_{k=1}^{\kappa(\nu)}p_k^{\alpha_k(\nu)} \tag{2.8}$$

its canonical representation.

Then, each random variable $\alpha(\nu,p)$ in (8) has a geometric probability distribution with a parameter $u=\dfrac{1}{p^s}$ $(0<u<1)$:

$$P_s\{\alpha(\nu,p)=a\}=u^a\cdot(1-u)=\left(\frac{1}{p^s}\right)^a\cdot\left(1-\frac{1}{p^s}\right),\quad a=0,1,2,\ldots \tag{2.9}$$

We have then,

$$E[\alpha(\nu,p]=\frac{1-u}{u}=p^s-1,\,Var[\alpha(\nu,p]=\frac{1-u}{u^2}=p^{2s}-p^s\,.$$

Variables $\alpha_k(\nu)=\alpha(\nu,p_k)$ are independent for all primes p_k $(k=1,2,\ldots,\kappa(\nu))$ as well as factors $p_k^{\alpha(\nu,p_k)}$ and $p_j^{\alpha(\nu,p_j)}$ for all $k\ne j$ in the canonical factorization $\nu=\prod_{p\in\mathbb{P}}p^{\alpha(\nu,p)}$.

Proof.

Denote $a\setminus b$ and $a\nmid b$ events $'a \text{ divides } b'$ and $'a \text{ does not divide } b'$, respectively.

We have:

$$P_s\{(p^k\setminus\nu)\cap(p^{k+1}\nmid\nu)\}=P_s\{p^k\setminus\nu\}-P_s\{p^{k+1}\setminus\nu\}\ \text{ since }\ \{p^{k+1}\setminus\nu\}\subset\{p^k\setminus\nu\}.$$

Notice that $P_s\left\{p^k \setminus \nu\right\} = P_s\left\{\nu \in p^k \cdot \mathbb{N}\right\} = \frac{1}{\zeta(s)} \cdot \sum_{m\in\mathbb{N}} \frac{1}{\left(p^k \cdot m\right)^s} = \left(\frac{1}{p^s}\right)^k \cdot \frac{1}{\zeta(s)} \cdot \sum_{m\in\mathbb{N}} \frac{1}{m^s} = \left(\frac{1}{p^s}\right)^k$.

Therefore,

$$P\left\{\left(p^{\alpha(\nu,p))} \setminus \nu\right) \cap \left(p \nmid \frac{\nu}{p^{\alpha(\nu,p)}}\right)\right\} = P\left\{\left(p^{\alpha(\nu,p)} \setminus \nu\right) \cap \left(p^{\alpha(\nu,p)+1} \nmid \nu\right)\right\} = \left(\frac{1}{p^s}\right)^{\alpha(\nu,p)} \cdot \left(1 - \frac{1}{p^s}\right).$$

Denote $E_m = C_{m,0}$ the event $\{m \setminus \nu\}$ ('m divides ν'). Then, for $m = m_1 \cdot m_2$ we have

$$P_{\zeta(s)}\left(E_m\right) = P_{\zeta(s)}\left(C_{m,0}\right) = \sum_{k\geq 11} \frac{\left(m \cdot k\right)^{-s}}{\zeta(s)} = \frac{1}{m^s} \sum_{k\geq 1} \frac{k^{-s}}{\zeta(s)} = \frac{1}{m^s} = \frac{1}{m_1^{\ s} \cdot m_2^s},$$

$$P_{\zeta(s)}\left(C_{m_i,0}\right) = \sum_{k\geq 11} \frac{\left(m_i \cdot k\right)^{-s}}{\zeta(s)} = \frac{1}{m_i^s} \sum_{k\geq 1} \frac{k^{-s}}{\zeta(s)} = \frac{1}{m_i^s} \ (i = 1, 2).$$

Similar,

$$\begin{aligned} P_{\zeta(s)}\left(C_{m_1 \cdot m_2,0}\right) &= \sum_{k\geq 11} \frac{\left(m_1 \cdot m_2 \cdot k\right)^{-s}}{\zeta(s)} = \sum_{k\geq 1} \frac{k^{-s}}{\zeta(s)} \cdot \frac{1}{m_1^s \cdot m_2^s} = \frac{1}{m_1^s \cdot m_2^s}. \\ &= P_{\zeta(s)}\left(C_{m_1,0}\right) \cdot P_{\zeta(s)}\left(C_{m_2,0}\right) \end{aligned}$$

If m_1 and m_2 are co-prime numbers, then $C_{m_1 \cdot m_2,0} = C_{m_1} \cap C_{m_2}$, that is $E_{m_1 \cdot m_2} = E_{m_1} \cap E_{m_2}$, and $P_s\left(E_{m_1} \cap E_{m_2}\right) = P_s\left(E_{m_1}\right) \cdot P_s\left(E_{m_2}\right)$, which holds true for any two different primes $m_1 = p_1$ and $m_2 = p_2$. This proves independence of $\alpha(p,\nu)$ for different primes p, as well as independence of factors $p_i^{\alpha(\nu,p_i)}$ and $p_j^{\alpha(\nu,p_i)}$ for all $i \neq j$ in the canonical factorization $\nu = \prod_{p\in\mathbb{P}} p^{\alpha(p,\nu)}$.

Q.E.D.

Theorem 2.2

A random variable ν with Zeta distribution $P_s\{\nu = n\} = \frac{n^{-s}}{\zeta(s)},\ s > 0,\ n \in \mathbb{N}$

represents a random walk $\{\nu(i) \mid 0 \leq i \leq \kappa(\nu)\}$ on a multiplicative semigroup $S(\mathbb{P}^*)$

generated by the extended set of primes $\mathbb{P}^* = \mathbb{P} \cup \{1\}$.

The walk on $\mathbb{P}^*$ is defined recursively as follows:

$$\begin{cases} \nu(1) = \nu(0)\cdot\eta(1), \text{ where } \nu(0)=1,\ \eta(1) = p_1^{\alpha_1(\nu)} \\ \nu(i+1) = \nu(i)\cdot\eta(i+1), \text{ where } \eta(i+1) = p_{i+1}^{\alpha_{i+1}(\nu)} (i = 0,1,2,\ldots,\kappa(\nu)-1) \end{cases} \tag{2.10}$$

Here random variables $\alpha_i = \alpha(\nu, p_i)$, due to Theorem 1, follow geometric

probability distributions (9) with parameters $u = \frac{1}{p_i^s}$ $(0 < u < 1)$, respectively.

The sequence $\{\nu(i) \mid 0 \le i \le \kappa(\nu)\}$ is a finite walk on $S(\mathbb{P}^*)$ with independent

multiplicative increments $\eta(i) = p_i^{\alpha_i(\nu)}$ such that $P\{\eta(i) = p_i^{a_i}\} = \left(\frac{1}{p_i^s}\right)^{a_i} \cdot \left(1 - \frac{1}{p_i^s}\right)$, and

$\kappa(\nu) \le \log_{p_{\min}} \nu = \frac{\ln \nu}{\ln p_{\min}}$, where $p_{\min}$ is the least prime number that divides ν.

Proof.

Formulas (1.7) and (1.9) imply: $\nu = \prod_{p\in\mathbb{P}} p^{\alpha(\nu,p)} = \left(\prod_{p:\,\alpha(\nu,p)=0} 1\right)\cdot\left(\prod_{p:\,\alpha(\nu,p)>0} p^{\alpha(\nu,p)}\right) = \prod_{k=1}^{\kappa(\nu)} p_i^{\alpha_i}$

Since $\xi(i) = p_i^{\alpha_i}$ and all $\alpha_i = \alpha(\nu, p_i)$, due to Theorem 1, are independent random

variables each with geometric distribution, we have $P\{\eta(i) = p_i^{a_i}\} = \left(\frac{1}{p_i^s}\right)^{a_i} \cdot \left(1 - \frac{1}{p_i^s}\right)$,

were $i = 1,2,\ldots,n,$ so that $\nu(n) = \prod_{i=1}^{n} \eta(i)$ for all n: $1 \le n \le \kappa(\nu)$ and $\nu(n) = \nu$ if $n = \kappa(\nu)$.

Thus, $P\{\nu = m\} = \prod_{i=1}^{\kappa(m)} \left(\frac{1}{p_i^s}\right)^{\alpha_i} \cdot \prod_{i=1}^{\infty}\left(1 - \frac{1}{p_i^s}\right) = \frac{1}{m^s}\cdot\frac{1}{\zeta(s)}$ since $m = \prod_{i=1}^{\kappa(m)} p_i^{\alpha_i}$.

Since $m = \prod_{i=1}^{\kappa(m)} p_i^{\alpha_i} \ge (p_{\min})^{\kappa(m)}$, where $p_{\min} \le p_i$ for all i: $1 \le i \le \kappa(m)$, we have:

$\kappa(m) \le \log_{p_{\min}} m$

Q.E.D.

Theorem 2.3

Let $h: R \to \{0,1\}$ be the Heaviside function $h(x) = \begin{cases} 1 \ \text{if} \, x > 0 \\ 0 \ \text{if} \, x \le 0 \end{cases}$,

$r(\nu_i) = \text{mod}(\nu, p_i), \vec{r}(\nu) = \left(r(\nu_i) \,|\, 1 \le i \le \pi\left(\sqrt{\nu}\right)\right)$ and

$\vec{\rho}(\nu) = \min\left(\vec{r}(\nu)\right) = \min_i \left(r(\nu_i) \,|\, 1 \le i \le \pi\left(\sqrt{\nu}\right)\right)$.

If a random variable ν has Zeta probability distribution and $\xi(n) = h\left(\vec{\rho}(n)\right)$, then for each $n \in \mathbb{N}$ the following statements hold true:

$$(1) \quad P_s\left\{\nu \in \mathbb{P} | \nu = n\right\} = P_s\left\{h\left(\vec{\rho}(\nu)\right) = 1\right\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p^s}\right) \tag{2.11}$$

$$(2) \quad P_s\left\{\xi(n) = \pi(n+1) - \pi(n) = 1 \,|\, \pi(1) = 1\right\} = P_s\left\{h(\rho(\nu) = 1\right\} = \prod_{p \le \sqrt{n+1}} \left(1 - \frac{1}{p^s}\right)$$

Proof.

Theorem 1 implies

$$P_s\left\{\nu \in \mathbb{P} \,|\, \nu = n\right\} = P_s\left\{\bigcap_{p \le \sqrt{\nu}} \left\{p \nmid \nu\right\} | \, \nu = n\right\} = \prod_{p \le \sqrt{\nu}} P_s\left\{p \nmid \nu \,|\, \nu = n\right\} = \prod_{i=1}^{\pi(\sqrt{n})} \left(1 - \frac{1}{p_i}\right)$$

Notice that the event $\left\{\bigcap_{p \le \sqrt{\nu}} \left\{p \nmid \nu\right\} | \, \nu = n\right\}$ can be expressed in the form of conditions

$$\left\{\bigcap_{p \le \sqrt{\nu}} \left\{\left[\text{mod}(\nu, p) > 0 \,|\, p \in \mathbb{P}\right]\right\}\right\} = \left\{\bigcap_{1 \le i \le \sqrt{\pi(\nu)}} \left\{r_i > 0\right\}\right\} = \left\{\min\left[r_i \,|\, 1 \le i \le \sqrt{\pi(\nu)}\right] > 0\right\}. \tag{2.12}$$

By using the Heaviside function $h(x) = \begin{cases} 1 \ \text{if} \, x > 0 \\ 0 \ \text{if} \, x \le 0 \end{cases}$, we can write the recursive equation (4.6) for $\pi(n)$ in the form:

$$\pi(n+1) = \pi(n) + h\left(\min_{p \le \sqrt{n}} \left\{\text{mod}(n, p) \,|\, p \in \mathbb{P}\right\}\right)$$

or, equivalently,

$$\pi(n+1)=\pi(n)+h\left(\min_{i\le\sqrt{n}}\left\{r_i \mid r_i=\operatorname{mod}(n,p_i)\right\}\right)=\pi(n)+h\left(\min(\vec{r}(n)\right) \qquad (2.13)$$

which controls the occurrence of prime numbers in the sequence of all integers $n=3,4,5,6,\ldots$ For a random number ν with Zeta probability distribution, vector of residuals $\vec{r}(\nu)=\left(r_1(\nu),r_2(\nu),\ldots,r_{\kappa(\nu)}(\nu)\right)$ is a vector with independent random components $r_k(\nu)=\operatorname{mod}(\nu,p_k)$ distributed within congruence classes $C_{p_k,r_k(\nu)}$ for all $k:\ 1\le k\le\pi(\nu)$. For ν to be prime, this is necessary and sufficient that ν should not be divisible by all of primes $p\le\sqrt{\nu}$, which means that $\min\left\{\vec{r}(\nu)\right\}=\min\left\{r_i(\nu)\,|\,1\le i\le\pi\left(\sqrt{\nu}\right)\right\}>0$. Denoting $\xi(n)=h\left(\vec{\rho}(n)\right)$ $(n=1,2,3,\ldots)$, we have:

$$P_s\left\{\xi(n)=\pi(n+1)-\pi(n)=1\,|\,\pi(1)=1\right\}=P\left\{h(\vec{\rho}(n)=1\right\}$$

$$=P_s\left\{\min\left\{(\vec{r}(n)>0\right\}\right\}=\frac{\frac{1}{\zeta(0)}\cdot\prod_{p\le\sqrt{n+1}}\left(1-\frac{1}{p^s}\right)}{\frac{1}{\zeta(0)}}=\prod_{p\le\sqrt{n+1}}\left(1-\frac{1}{p^s}\right) \qquad (2.14)$$

Therefore, by letting $s\to1$, we obtain

$$P_s\left\{\xi(n)=1\,|\,\pi(0)=0\right\}\to\prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right) \qquad (2.15)$$

Probability of a random ν to be a prime number in the interval $[2,\ n]$ for all $n\ge5$ is given by the formulas:

$$P\left\{\nu\in\mathbb{P}|\nu\le n\right\}=P\left\{h\left(\vec{\rho}(\nu)\right)=1\right\}=\prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right),$$

$$P\left\{\xi(n)=1\,|\min\left(\min(\rho_i\,|\,1\le i\le\pi(\sqrt{n})\right)>0\right\}=\prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right) \qquad (2.16)$$

Examples.

1) $\nu=108=1\cdot2^2\cdot3^3\cdot5^0\cdot7^0\cdots$ with $\alpha(108,p)=0$ for all $p>3$.

We have: $\alpha(108,2)=2, \alpha(108,3)=3; \kappa(108)=2$

2) ν=110=$2\cdot 3^0\cdot 5\cdot 7^0\cdot 11\cdot 13^0\cdot 17^0\cdots$ with $\alpha(110,p)=0$ for $p=3, 7$, and all $p>11$

We have: $\alpha(110,2)=1, \alpha(110,5)=1, \alpha(110,11)=1;\ \kappa(110)=3$.

In the above setting, the number $108=\prod_{i=0}^{\infty}\xi(i)$ in example 1) represents the path:

$$1\to 2^2\to 3^3\to 5^0\to 7^0\to\cdots$$

The number $110=\prod_{i=0}^{\infty}\xi(i)$ in example 2) represents the path:

$$1\to 2\to 3^0\to 5\to 7^0\to 11\to 13^0\to 17^0\to\ldots$$

By setting $P\left\{\xi(j)=p_j^{\alpha_j}\right\}=\left(\frac{1}{p_j^s}\right)^{\alpha_j}$ for all $p_j\in\mathbb{P}$, we can calculate probability $P\{\nu=n\}$ of any given value $n\in\mathbb{N}$.

In example 1):

$$P_s\{\nu=108\}=\frac{1}{2^{2s}}\cdot\left(1-\frac{1}{2^s}\right)\cdot\frac{1}{3^{3s}}\cdot\left(1-\frac{1}{3^s}\right)\cdot\left(1-\frac{1}{7^s}\right)\cdot\left(1-\frac{1}{11^s}\right)\cdots\left(1-\frac{1}{p_j^s}\right)\cdots$$

$$=\frac{1}{2^{2s}\cdot 3^{3s}}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{1}{108^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$$

In example 2):

$$P_s\{\nu=110\}=\frac{1}{2^s}\cdot\left(1-\frac{1}{2^s}\right)\cdot\left(1-\frac{1}{3^s}\right)\cdot\frac{1}{5^s}\cdot\left(1-\frac{1}{5^s}\right)\cdot\left(1-\frac{1}{7^s}\right)\cdot\frac{1}{11^s}\cdot\left(1-\frac{1}{11^s}\right)\cdot\left(1-\frac{1}{13^s}\right)\cdots\left(1-\frac{1}{p_j^s}\right)\cdots$$

$$=\frac{1}{2^s\cdot 5^s\cdot 11^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)=\frac{1}{110^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$$

Notice that, in general, in the formal expression $P_s\{\nu=n\}=\frac{1}{n^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right)$ the product involves a set of all prime numbers. In the above expressions the ‘probability’ $P_s\{\nu = n\}$ depends on a parameter s:

$$P_s\{\nu = n\} = \frac{1}{n^s}\cdot\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j^s}\right) = \frac{1}{n^s\cdot\zeta(s)},\ \ n\in\mathbb{N}, s>1 \tag{2.17}$$

The probability distribution (5.6) is called Riemann Zeta probability distribution.

To cope with the divergence of the infinite product $\prod_{j=1}^{\infty}\left(1-\frac{1}{p_j}\right) = \prod_{p\in\mathbb{P}}\left(1-\frac{1}{p}\right) = \zeta(1)$,

we consider $\zeta(s) = \sum_{n\geq 1}\frac{1}{n^s}$ where $s>1$, and define the probability P_s as a function of a parameter s. Meanwhile, there is another way to cope with divergence of $\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s}$ for $s\leq 1$. We can do so by introducing a sequence of *incomplete* (or *partial*) *Riemann Zeta* functions. We define the *incomplete product Zeta function* $\zeta_{\mathbb{P}_N}(s)$ as a partial product in the multiplicative presentation of $\zeta(s)$ for $s\geq 1$:

$$\zeta_{\mathbb{P}_N}(s) = \prod_{p\leq N}\frac{1}{1-\frac{1}{p^s}} \tag{2.18}$$

Remark 2.1.

Since $\frac{1}{1-\frac{1}{p^s}} = \sum_{k=0}^{\infty}\left(\frac{1}{p^s}\right)^k$, we have a convergent additive partial presentation of $\zeta(s)$:

$$\zeta_{\mathbb{P}_N}(s) = \prod_{p\leq N}\left[\sum_{k=0}^{\infty}p^{-sk}\right] = \sum_{n\in S(\mathbb{P}_N)} n^{-s}. \tag{2.19}$$

Here $S(\mathbb{P}_N)$ is a multiplicative semigroup of all integers generated by $\mathbb{P}_N{}^* =$ $\mathbb{P}_N\cup\{1\}$, where $\mathbb{P}_N = \{p \mid p\leq N, p\in\mathbb{P}\}$. Notice that $S(\mathbb{P}_N)$is an infinite set generated by a finite set $\mathbb{P}_N{}^*$. Then, we consider the corresponding probability distribution $P_{s,N}$, $s>1$, given by the formula:

$$P_{s,N}\{\nu = n\} = \frac{1}{n^s \cdot \zeta_{\mathbb{P}_N}(s)},\ n \in S(\mathbb{P}_N),\ s > 0,\ N \in \mathbb{N} \qquad (2.20)$$

Since $\zeta_{\mathbb{P}_N}(s) = \prod_{p \le N}\left(1 - \frac{1}{p^s}\right)^{-1} = \sum_{n \in S(\mathbb{P}_N)} \frac{1}{n^s}$, we have $\sum_{n \in S(\mathbb{P}_N)} P_{s,N}\{\nu = n\} = 1$.

The probability $P_{s,N}$ of ν to be a prime number in the set of numbers $S(\mathbb{P}_N)$ (generated by primes not exceeding N) can be calculated by the formula:

$$P_{s,N}\{\nu \text{ is prime} | \nu \in S(\mathbb{P}_N)\} = \frac{\sum_{p \in \mathbb{P}_N} p^{-s}}{\zeta_{\mathbb{P}_N}(s)} = \frac{\sum_{p \le N} p^{-s}}{\prod_{p \le N} \frac{1}{1 - p^{-s}}} = \left(\sum_{p \le N} p^{-s}\right) \cdot \prod_{p \le N}\left(1 - \frac{1}{p^s}\right) \qquad (2.21)$$

The convergence of the infinite series $\zeta_{\mathbb{P}_N}(z) = \sum_{n \in S(\mathbb{P}_N)} n^{-s}$ is guaranteed by (2.18) and (2.19). In general, from the probabilistic point of view, every finite path on the monoid set $S(\mathbb{P}^*) = \mathbb{N}$ can be identified with a randomly chosen natural number ν by assuming that it has a probability distribution

$$P\{\nu = n\},\ n \in \mathbb{N},\ \text{such that } \sum_{n=1}^{\infty} P\{\nu = n\} = 1 .$$

3. Asymptotics of a generalized Bernoulli process and the Cramér's model of prime numbers distribution

Definition 3.1

A sequence of $\{0,1\}$-valued random variables $(\xi_k)_{k\in\mathbb{N}}$ defined on probability space $(\Omega,\mathcal{F},P)$ which terms are not in general independent and dentically distributed we call *a generalized Bernoulli process*. We have:

$$P\{\xi_k=1\}=P_k, P\{\xi_k=0\}=Q_k,\ P_k+Q_k=1,\ k\in\mathbb{N}.$$

Probabilitstic approach to distribution of prime numbers in $\mathbb{N}$ is addresed in the Harald Cramér's model [2,3]. The Cramér's model describes the occurrence of prime numbers as a special case of a Bernoulli process given by a *sequence of independent Bernoulli variables* $(\xi(n)\,|\,n\in\mathbb{N})$, where $\xi(n)=\xi(\nu_n)$, with probabilities

$$P\{\nu_n\in\mathbb{P}\}=\frac{1}{\ln n},\ \ P\{\nu_n\notin\mathbb{P}\}=1-\frac{1}{\ln n},$$

or equivalently,

$$P\{\xi(n)=1\}=\frac{1}{\ln n},\ \ P\{\xi(n)=0\}=1-\frac{1}{\ln n}\ ,\ \text{where } n\geq 2.$$

The above formulas, due to the Merten's 1st and 2nd theorems [2, p.15], have the asymptotic expression:

$$P\{\xi(n)=1\}=\prod_{p\leq\sqrt{n}}\left(1-\frac{1}{p}\right)\approx\frac{e^{-\gamma}}{\frac{1}{2}\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right]=\frac{c}{\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right], \tag{3.1}$$

where $c=\frac{2}{e^{\gamma}}\approx 1.122918968\cdot$

We consider all values of $n>N$ in (17) by choosing arbitrary large natural N.

As we pointed above, the Cramér's assumption about *independence of terms* in the sequence $\left(\xi(n) \mid n=1,2,\ldots\right)$ is not quite accurate. The more realistic and adequate approach would be to consider the sequence of consecutive primes represented by $\left(v_n\right)_{n\in\mathbb{N}}$ and $\left(\xi(n) \mid n=1,2,\ldots\right)$, respectively, as *stochastic predictable sequences of dependent random variables.*

Actually the sequence of random variables in the Cramér's model is *asymptotically Bernoullian* (and asyptotically pairwise independent) in a sense of Definition 3.1 given below. Meanwhile, the demand for idependence of terms in the sequence $\left(v_n\right)_{n\in\mathbb{N}}$ and in $\left(\xi(n)\right)_{n\in\mathbb{N}}$ could be be relaxed in the Cramér's model, due to the version of the *Central Limit Theorem (CLT) for dependent random variables*.

This version of CLT tracks back to the S.N. Bernstein's ideas. In rather general terms its generalization is proved in [19 - 21] for random walks on differntiable maniforlds and Lie groups. We provide here the formulation of this theorem given in [17] as the most adequate for the goals of this article.

First, we discuss, following M. Loèv [19], *asymptotic behavior of a generalized Bernoulli process.* We have for ξ_k mathematical expectation $E\left\{\xi_k\right\}=P_k$ and variance $V\left\{\xi_k\right\}=P_k\cdot Q_k$.

Let's denote $X_n=\frac{1}{n}\sum_{k=1}^{n}\xi_n$. Then $E\left\{X\right\}_n=\frac{1}{n}\sum_{k=1}^{n}P_k$. Since $\left(\xi_k\right)^2=\xi_k$, we have

$E\left\{\left(\xi_k\right)^2\right\}=E\left\{\xi_k\right\}, E\left\{\xi_k\cdot\xi_l\right\}=P\left\{\xi_k\cdot\xi_l=1\right\}=P_{kl}$. Then,

$\left(E\left\{X_n\right\}\right)^2=\frac{1}{n^2}\left(\sum_{k=1}^{n}P_k^2+2\sum_{1\le k<l<]\le n}P_k\cdot P_l\right)$ and $E\left\{\left(X_n\right)^2\right\}=\frac{1}{n^2}\left(\sum_{k=1}^{n}P_k+2\sum_{k<l}P_{kl}\right)$.

This implies:

$$V\left\{X_n\right\}=E\left\{(X_n)^2\right\}-\left(E\left\{X_n\right\}\right)^2=\frac{1}{n^2}\left(\sum_{k=1}^{n}P_kQ_k+2\sum_{1\le k<l\le n}(P_{kl}-P_k\cdot P_l)\right)=\frac{1}{n^2}\sum_{k=1}^{n}V(\xi_k)+D_n\ ,$$

where $D_n = \frac{2}{n^2}\sum_{1\le k<l\le n}(P_{kl} - P_k \cdot P_l) = \frac{n(n-1)}{2n^2}\left(\frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_{kl} - \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_k \cdot P_l\right)$.

If terms in $\left(\xi(n)\right)_{n\in\mathbb{N}}$ are pairwise independent, then

$P_{kl} = E\{\xi_k \cdot \xi_l\} = E\{\xi_k\}\cdot E\{\xi_l\} = P_k \cdot P_l$ and $D_n = 0$ which implies $V\{X_n\} = \frac{1}{n^2}\sum_{k=1}^{n} V\{\xi_k\}$.

Thus, D_n can be viewed as a *cummulative measure of pairwise independence* of terms in a Bernoulli process $\left(\xi(n)\right)_{n\in\mathbb{N}}$. Denote:

$$\overline{P}_1(n) = \frac{1}{n}\sum_{k=1}^{n} P_k \text{ and } \overline{P}_2(n) = \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_{kl}.$$

Notice that

$$D_n = \frac{n-1}{2n}\left(\overline{P}_2 - \overline{P}_{1,2}\right) \text{ where } \overline{P}_{12} = \frac{2}{n(n-1)}\sum_{1\le k<l\le n} P_k \cdot P_l.$$

We consider below a slightly different measure d_n that shows how close a Bernoulli process $\left(\xi(n)\right)_{n\in\mathbb{N}}$ is to a classical Bernoulli ssequence of independent equally distributed random variables.

Then, $E\left\{\left(X_n\right)^2\right\} = \frac{1}{n^2}\left(\sum_{k=1}^{n} P_k + 2\sum_{k<l} P_{kl}\right) = \frac{\overline{P}_1 - \overline{P}_2}{n} + \overline{P}_2$ and $\left(E\{X_n\}\right)^2 = \left(\overline{P}_1\right)^2$

Since $V\{X_n\} = E\{(X_n)^2\} - \left(E\{X_n\}\right)^2 = \frac{\overline{P}_1 - \overline{P}_2}{n} + \overline{P}_2 - \left(\overline{P}_1\right)^2$, we have

$$V\left(X_n\right) = \frac{\overline{P}_1 - \overline{P}_2}{n} + \overline{P}_2 - \left(\overline{P}_1\right)^2 = \frac{\overline{P}_1 - \overline{P}_2}{n} + d_n \tag{3.2}$$

where $d_n = \overline{P}_2 - \left(\overline{P}_1\right)^2$.

In the classical Bernoulli scheme witn independent identically distributed terms $\left(\xi_k\right)_{k\in\mathbb{N}}$, we have $P_{kl} = E\{\xi_k \cdot \xi_l\} = P\{\xi_k \cdot \xi_l = 1\} = P\{\xi_k = 1\}\cdot P\{\xi_l = 1\} = P_k \cdot P_l = P^2$, due to independence and equal distribution of terms in the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$, so that $d_n = \overline{P}_2 - \left(\overline{P}_1\right)^2 = P^2 - P^2 = 0$. This implies $d_n = 0$and $V\{X_n\} = \frac{1}{n^2}\sum_{k=1}^{n} V\{\xi_k\}$.

This means that the value of d_n is a measure of a deviation of the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ from a classical Bernoulli scheme.

Definition 3.1

We call a sequence $\left(\xi_n\right)_{n\in\mathbb{N}}$ of $\{0,1\}$-valued random variables defined on probability space $(\Omega,\mathcal{F},P)$ *asymptotically pairwise Bernoullian* if $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\to 0$ as $N\to\infty$. This means that for sufficiently large N variables ξ_k,ξ_l are asymptotically independent for all $l>k>N$.

Lemma 3.1

For *asymptotically Bernoullian sequence* $\left(\xi_n\right)_{n\in\mathbb{N}}$ we have $D_n\to 0$ so that

$$\left|V\{X_n\}-\frac{1}{n^2}\sum_{k=1}^{n}V\{\xi_k\}\right|\to 0 \text{ as } n\to\infty .$$

Proof.

Due to (5), $V\{X_n\}-\frac{1}{n^2}\sum\limits_{k=1}^{n}V\{\xi_k\}=D_n$.

Since $D_n=\frac{2}{n^2}\sum\limits_{k<l\le n}\left(P_{kl}-P_k\cdot P_l\right)$, and $\left|\sum\limits_{k<l\le n}\left(P_{kl}-P_k\cdot P_l\right)\right|\le\frac{n(n-1)}{2}\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|$,

we have $\left|D_n\right|\le\frac{2}{n^2}\cdot\frac{n(n-1)}{2}\cdot\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\to 0$.

This implies $\left|V\{X_n\}-\frac{1}{n^2}\sum\limits_{k=1}^{n}V\{\xi_k\}\right|\to 0$.

Q.E.D.

Keeping in mind approximation (17), we restrict the sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ by considering its 'tail' $\left(\xi_k\right)_{k>N}$ of the original seqience for sufficiently large N.

Theorem 3.1

The sequence $\left(\xi_k\right)_{k\in\mathbb{N}}$ in the modified Cramér's model is *asymptotically pairwise Bernoullian,* that is $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|=O\left(\frac{1}{\ln N}\right)$, where $P\left\{\xi_k=1\right\}=P_k$, $P\left\{\xi_k\cdot\xi_l=1\right\}=P_{kl}$,

and

$$\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\frac{1}{\ln N}\right) \quad \text{as } D_n=O\left(\frac{1}{\ln N}\right) \text{ for all } n>N. \qquad (3.3)$$

Proof.

Indeed, $P\left\{\xi_k=1\right\}=P_k$, $P\left\{\xi_k\cdot\xi_l=1\right\}=P_{kl}$. Then, since $\left|P_{kl}-P_k\cdot P_l\right|<P_k\le\frac{1}{\ln N}$ for all $N<k<l\le n$, we have $\max\limits_{N<k<l}\left|P_{kl}-P_k\cdot P_l\right|\le\frac{c}{\ln N}\to 0$ and $D_n=O\left(\frac{1}{\ln N}\right)$ as $N\to\infty$.

This implies $\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\frac{1}{\ln N}\right)$ for all $n>N$ (3.4)

Q.E.D.

In the Cramér's model $\hat{\pi}_N\left(n\right)=\sum_{k=N}^{N+n}\xi(k)$ represents the number of primes among n terrms in the interval $(N,N+n]$ of the sequence and $\frac{\hat{\pi}_N(n)}{n}=\frac{1}{n}\sum_{k=N}^{N+n}\xi(k)=\frac{\hat{\pi}_N(n)}{n}$ is a relative freqiency of primes for these terms. predicted by the imprved model.

In the Table 4 below, we demonstrate how well $E\left\{\frac{\hat{\pi}(n)}{n}\right\}$ approximates relative frequencies of primes $\frac{\pi(n)}{n}$ in the improved Cramér's model $\left(\xi_k\right)_{k\ge 3}$ as n increses from 10^1 to 10^9.

Table 3.1. *Comparison of probabilities y* $P\{v \in \mathbb{P} | v \le n\}$

and frequencies $\frac{\pi(n)}{n}$

Natural n	$P\{v \text{ is prime} \mid v \le n\} = \prod_{p \le \sqrt{n}} \left(1 - \frac{1}{p}\right)$	Relative frequency $\frac{\pi(n)}{n}$ of primes in intervals $[1, n]$
10^1	0.33333333	0.400000 00
10^2	0.22857143	0.25000000
10^3	0.15285215	0.16800000
10^4	0.12031729	0.12290000
10^5	0.09621491	0.09592000
10^6	0.08096526	0.07849800
10^7	0.06957939	0.06645790
10^8	0.06088469	0.05761455
10^9	0.05416682	0.05084753

Consider now the *Generalized Law of Large Numbers* for a general Bernoulli process as it stated in[18] and apply it then to the improved Cramér's model $(\xi_k)_{k \in \mathbb{N}}$.

Theorem 3.2

Let $\xi(k) = \begin{cases} 1 \text{ if } k \in \mathbb{P} \\ 0 \text{ otherwise} \end{cases}$ and $\frac{\hat{\pi}_N(n)}{n} = \frac{1}{n} \sum_{k=N+1}^{N+n} \xi(k)$ be a relative freqiency of primes the interval $[N, N+n]$. Then, the *Generalized Law of Large Numbers* holds true:

$$P\left\{\left|\frac{\hat{\pi}_N(n)}{n}-E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}\right|>\varepsilon\right\}\to 0 \text{ as } N,n\to\infty.$$

(21)

If $d_{n.N}=\frac{2}{n(n-1)}\sum\limits_{N\le k<l\le N+n}P\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\}-\left(\frac{1}{n}\sum\limits_{k=N}^{N+n}P\{k\in\mathbb{P}\}\right)^2=O\left(\frac{1}{n}\right)$,

then the *Generalized Strong Law of Large Numbers* holds true:

$$P\left\{\left|\frac{\hat{\pi}_N(n)}{n}-E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}\right|\to 0\right\}=1 \text{ as } N,n\to\infty \tag{3.5}$$

Proof.

Due to [25], we apply the following *Propositions:*

1. The Generalized Bernoulli Theorem that for every $\varepsilon>0$: $P\{|X_n-E\{X_n\}|>\varepsilon\}\to 0$

holds true for a Bernoulli process $(\xi_k)_{k\in\mathbb{N}}$ if and only if $d_n=\overline{P}_2-(\overline{P}_1)^2\to 0$ as $n\to\infty$.

2. The Generalized Strong form of Bernoulli Theorem that $P\{|X_n-E\{X_n\}|\to 0\}\to 1$

holds true if $d_n=O\left(\frac{1}{n}\right)$.

We show here that these propositions asymptotically hold true for tails $(\xi_k)_{k\ge N}$ of

the Cramér's model $(\xi_k)_{k\ge 3}$. For a tail $(\xi_k)_{k\ge N}$ of the Cramér's model we have:

$$\overline{P}_{1,N}(n)=E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}=\frac{1}{n}\sum_{k=N+1}^{N+n}E\{\xi(k)\}=\frac{1}{n}\sum_{k=N+1}^{N+n}P\{k\in\mathbb{P}\}=\frac{1}{n}\sum_{k=N+1}^{N+n}\frac{1}{\ln k}\sim\int_N^{N+n}\frac{dt}{\ln t}=Li(N+n)-Li(N)$$

$$\overline{P}_{2,N}(n)=\frac{2}{n(n-1)}\sum_{N\le k<l\le N+n}E\{\xi(k)\cdot\xi(l)\}=\frac{2}{n(n-1)}\sum_{N\le k<l\le N+n}P\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\}\cdot$$

Then, $d_{n.N}=\frac{2}{n(n-1)}\sum\limits_{N\le k<l\le N+n}P\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\}-\left(\frac{1}{n}\sum\limits_{k=N}^{N+n}P\{k\in\mathbb{P}\}\right)^2$

Notice that $d_{n.N}\le\max\limits_{N\le k<l\le N+n}\left(P\{(k\in\mathbb{P})\cap(l\in\mathbb{P})\}\right)<\frac{1}{\ln N}$ implies $d_{n,N}\to 0$ as $n,N\to\infty$.

This implies $d_{n,N} \to 0$ as $N, n \to \infty$

Then, $P\left\{\left|\frac{\hat{\pi}_N(n)}{n} - E\left\{\frac{\hat{\pi}_N(n)}{n}\right\}\right| > \varepsilon\right\} = O\left(\frac{1}{\ln N}\right)$ as $n, N \to \infty$ and (9) holds true.

If, in addition $d_{n,N} = O\left(\frac{1}{n}\right)$, as $n, N \to \infty$, then (10) holds true.

Q.E.D.

4. Asymptotic Distribution of Residuals

Notice that the vector function $r(n)=\mathrm{mod}\left(n,\vec{p}(n)\right)$ is periodic with a period $T=\prod_{p\le n} p$ since $\mathrm{mod}(T,p)=0$ for any $p\le n$. Due to the Chinese Remainder Theorem (CRT) [22, p.101], a solution x to the system of equations $\mathrm{mod}(x,p_i)=r_i\ \left(1\le i\le m\right)$ exists, and if x is a solution to the system, then $y=x+T$ is also a solution to the same system. Considering the ring of all integers $\mathbb{Z}$, we write $\mathbb{Z}_m=\mathbb{Z}/\left(m\cdot\mathbb{Z}\right)$.

Here $\mathbb{Z}_m$ consists of m congruence classes: $\mathbb{Z}_m=\left\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\right\}$ modulo m, also called *residue classes,* denoted as $\left[0\right]_m,\left[1\right]_m,\ldots,\left[m-1\right]_m$ with the addition and multiplication rules expressed as

$$\left[k\right]_m+\left[l\right]_m=\left[\mathrm{mod}(k+l,m)\right]_m \text{ and } \left[k\right]_m\cdot\left[l\right]_m=\left[\mathrm{mod}(k\cdot l,m)\right]_m,$$

respectively. For any prime number $p\in\mathbb{P}$, set $\mathbb{Z}_p$ of congruence classes modulo p is a *finite abelian group* $G_p=\mathbb{Z}_p=\mathbb{Z}/\left(p\cdot\mathbb{Z}\right)$, *of order* p.

Consider a random sequence $\omega=\left(\eta_1,\eta_2,\ldots,\eta_n\right)$ where $\eta_i\in G_{p_i}\,(i=1,2,\ldots,n)$ such that random variables $\eta_1,\eta_2,\ldots,\eta_n$ are mutually independent and we can always find the minimal solution to $\mathrm{mod}(x,p_i)=r_i\ \left(1\le i\le m\right)$ among all solutions.

For example, given $\vec{p}=(5,11,17,23,29)$ and $\vec{r}=(0,8,13,7,1)$, the system $\mathrm{mod}(x,p_i)=r_i\ \left(1\le i\le 5\right)$ has the minimal solution $x=30$. One of other possible solutions, for instance, is $x=623675$.

We are interested in probability measures on the direct product $\mathcal{G} = \prod_{p \in \mathbb{P}} G_p$ such that each non-trivial probability distribution is supported by a finite number of components in $\mathcal{G}$.

For a random sequence $\omega = (\eta_1, \ldots, \eta_n)$ of mutually independent random variables η_i $(i = 1, 2, \ldots, n)$ with distributions $P\{\eta_i = r \mid r \in G_{p_i}\} = q_r^{(i)}$ on G_{p_i}, we have

$$P\left\{\eta_i \in B \subseteq G_{p_i}\right\} = \sum_{r \in B} q_r^{(i)}, \quad \sum_{r=0}^{p_i - 1} q_r^{(i)} = 1 \ (i = 1, 2, \ldots, n) \ .$$

and

$$P\left\{\omega \in \prod_{i=1}^{n} B_i\right\} = \prod_{i=1}^{n} P\{\eta_i \in B_i\} \text{ for any } B_i \subset G_{p_i} \qquad (4.1)$$

Further, we use the following notation: $B - r \equiv \left\{ s \in G_p \middle| s + r \in B, r \in G_p \right\}$ and for every probability distribution P on G_p define the 'shifted' measure $\theta_r P(B) = P(B - r)$. Obviously the shifted measure $\theta_r P$ is a probability measure on subsets of a finite set $G_p : \theta_r P(G_p) = P(G_p - r) = 1$ because $G_p - r = G_p$ for any $r \in G_p$ since G_p is a group. Due to CRT, there exist one-to one correspondence between finite sequences of residues $(r_1, r_2, \ldots, r_n)$ and positive integers $n = \prod_{i=1}^{k} p_i^{\alpha_i}$ such that $\mathrm{mod}(n, p_i) = r_i$ $(i = 1, 2, \ldots, k)$. If $\mathrm{mod}(m, p_i) = s_i$ for some number m, then $\mathrm{mod}(n + m, p_i) = \mathrm{mod}(r_i + s_i, p_i)$. Consider two independent random integers ν and μ with probability measures P^{ν} and P^{μ}, and their residuals $[\nu]_p, [\mu]_p$ modulo p, respectively. We are interested in probability distribution $P^{[\nu+\mu]_p}$ of the sum $[\nu]_p + [\mu]_p = [\nu + \mu]_p$. For any set $B \subset G_p$ we have

$$P\left\{\left[\nu+\mu\right]_p \in B\right\} = \sum_{(r+s)\in B} P\left\{\left[\nu\right]_p = r\right\}\cdot P\left\{\left[\mu\right]_p = s\right\} = \sum_{t\in B} P\left\{\left[\nu\right]_p = t-s\right\}\cdot P\left\{\left[\mu\right]_p = s\right\}$$

and we denote $P^{[\nu+\mu]_p}(B) = P\left\{\left[\nu+\mu\right]_p \in B\right\}$ as

$$P^{[\nu+\mu]_p}(B) = P^{[\nu]_p} * P^{[\mu]_p}\left(B\right), \tag{4.2}$$

so that $P^{\nu+\mu}(B) = P^{\nu} * P^{\mu}\left(B\right) = \sum_{t\in B} P\left\{\left[\nu\right]_p = t-s\right\}\cdot P\left\{\left[\mu\right]_p = s\right\}$

The measure $P^{\nu+\mu}(B) = P^{\nu} * P^{\mu}\left(B\right)$ is called a *convolution* of measures P^{ν} and P^{μ}.
One of interesting questions is an asymptotic distribution of sums of independent random integers $\nu^{(n)} = \nu_1 + \nu_2 + \cdots + \nu_n$ and their corresponding residuals

$$\left[\nu^{(n)}\right]_p = \left[\nu_1\right]_p + \left[\nu_2\right]_p + \cdots + \left[\nu_n\right]_p$$

which are also sums of independent random variables $\left[\nu_i\right]_p$ $(i = 1, 2, \ldots, n)$.

The answer to the question about the limit distribution of $\nu^{(n)}$ depends in general on the distributions of the terms ν_i in the sum. Meanwhile the limit behavior of residuals $\left[\nu^{(n)}\right]_p$ does not depend (under very simple and natural conditions) on the distribution of each term $\left[\nu_i\right]_p$. In what follows we use the well-known general facts from Probability Theory regarding characteristic functions of probability distributions and their convolutions.
Let P^{ξ} be a probability measure defined on all finite subsets of $\mathbb{N}$. This means that for every $n \in \mathbb{N}$ there exists $P^{\xi}(n) = P\left\{\xi = n\right\} \geq 0$ such that $\sum_{n\in} P^{\xi}(n) = 1$.

Characteristic function Φ^{ξ} is defined by the formula

$$\Phi^{\xi}(t) = Ee^{it\cdot\xi} = \sum_{n\in\mathbb{N}} e^{it\cdot n} \cdot P^{\xi}(n).$$

For a finite abelian additive group $G_p = \mathbb{Z}_p$ we consider a homomorphism χ of G_p into multiplicative group C^* of complex numbers $\chi : G_p \to C^*$.

A homomorphism $\chi : G_p \to C^*$ is also called a *character*.

Since any element $[k]_p \in G_p$ $(k = 0,1,\ldots,p-1)$ has order p, that is $p \cdot [k]_p = [0]_p$,

we have $1 = \chi([0]_p) = \chi(p \cdot [k]_p) = (\chi([k]_p))^p$. This means that any character value $\chi([k]_p)$ is a p-th root of unity.

We can define p such character values: $\chi_r([k]_p) = e^{\frac{2\pi i}{p}(r \cdot k)}$ $(r = 0,1,2,\ldots,p-1)$.

Denote $\chi_{rk} = e^{\frac{2\pi i}{p}(r \cdot k)}$ $(r,k = 0,1,2,\ldots,p-1)$. Character $\chi_0([k]_p) = 1$ for all $k = 0,1,\ldots,p-1$, and χ_0 is called a *principal character*.

Consider a square matrix $\chi = [\chi_{rk}]$ $(0 \le r,k \le p-1)$ of size p. All characters are orthogonal to each other in terms of scalar products of rows of matrix χ :

$$\langle \chi_r, \chi_s \rangle = \sum_{t=0}^{p-1} \chi_{rt} \cdot \bar{\chi}_{st} = \sum_{t=0}^{p-1} e^{\frac{2\pi i}{p}(r \cdot t)} \cdot e^{\frac{-2\pi i}{p}(s \cdot t)} = \sum_{t=0}^{p-1} e^{\frac{2\pi i}{p}((r-s)\cdot t)} = \frac{1 - e^{2\pi i \cdot (r-s)}}{1 - e^{\frac{2\pi i \cdot (r-s)}{p}}} = \begin{cases} p, & \text{if } r = s \\ 0, & \text{if } r \ne s \end{cases}$$

Characteristic function $\Phi^{[\xi]_p}$ for residual $[\xi]_p$ is given by the formula

$$\Phi^{[\xi]_p}(r) = Ee^{i\chi_r([\xi]_p)} = \sum_{k=0}^{p-1} P^{[\xi]}(k) e^{\frac{2\pi i}{p}(r \cdot k)} = \sum_{k=0}^{p-1} \chi_{rk} \cdot P^{[\xi]}(k) = \left[\chi \cdot P^{[\xi]}\right](r)$$

Since the matrix $\chi = [\chi_{rk}]_p$ $(0 \le r,k \le p-1)$ is orthogonal, the inverse matrix χ^{-1} exists and the probability distribution $P^{[\xi]_p}$ can be uniquely recovered as $P^{[\xi]_p} = \chi^{-1} \cdot \Phi^{[\xi]_p}$ given its characteristic function $\Phi^{[\xi]_p}$.

There is one-to-one correspondence between finite probability distributions and the corresponding characteristic functions.

A probability distribution $P^{\xi}(k)\ (k=1,2,\ldots,n)$ defined on a finite set $X=\{x_1,x_2,\ldots,x_n\}$ can be identified with the n-dimensional vector $P^{\xi}=(p_1,p_2,\ldots,p_n)$ where $p_k=P\{\xi=k\}$, $1\le k\le n$.

If we have a sequence of probability distributions $P^{\xi_m}\ (m=1,2,\ldots)$ such that $P^{\xi_m}\to P$ in a sense of vector convergence in n-dimensional vector space to probability measure P on X , then we can expect the convergence for the sequences of corresponding characteristic functions: $\Phi^{\xi_m}\to\Phi$, where Φ is a characteristic function of some limit random variable ξ_∞ on X , and vice versa. One of the most important properties of characteristic functions is that for any two independent random variables ξ_1,ξ_2 we have $\Phi^{\xi_1+\xi_2}=\Phi^{\xi_1}\cdot\Phi^{\xi_2}$, so that $\Phi^{\sum_{i=1}^{n}\xi_i}=\prod_{i=1}^{n}\Phi^{\xi_i}$ for independent $\xi_1,\xi_2,\ldots,\xi_n$.

Theorem 4.1

For any random integers ν its residual $[\nu]_p$ for a prime $p\in\mathbb{P}$ has a characteristic function $\Phi^{[\nu]_p}$ such that $\Phi^{[\nu]_p}(0)=1$ and $\left|\Phi^{[\nu]_p}(r)\right|<1,$ if $0<r\le p-1$.

Proof.

If a random integer λ is such that $[\lambda]_p$ has a uniform distribution on G_p , that is $P\{[\lambda]_p=k\}=\frac{1}{p}$ for all $k=0,1,\ldots,p-1$, then $\Phi^{[\lambda]_p}(r)=\begin{cases}1, & \text{if } r=0\\ 0, & \text{if } r\ne 0\end{cases}$

We prove this by the direct calculations:

$$\Phi^{[\lambda]_p}(r)=\sum_{k=0}^{p-1}\chi_{rk}\cdot P^{[\lambda]_p}(k)=\sum_{k=0}^{p-1}\chi_{rk}\cdot\frac{1}{p}=\frac{1}{p}\langle\chi_r,\chi_0\rangle=\begin{cases}1, & r=0\\ 0, & r\ne 0\end{cases}$$

We have $\Phi^{[\nu_i]_p}(r)=\sum_{k=0}^{p-1}\chi_{rk}\cdot P^{[\nu_i]}(k)=\left[\chi\cdot P^{[\nu_i]}\right](r)$. This implies $\left|\Phi^{\nu_i}(r)\right|\leq 1$.

We have $\Phi^{\nu_i}(0)=1$. Assume that there exist $r\neq 0 \bmod p$ such that $\Phi^{\nu_i}(r)=1$.

Then, $\Phi^{[\nu]_p}(r)=\sum_{k=0}^{p-1}P^{[\nu]}(k)e^{\frac{2\pi i}{p}(r\cdot k)}=1$ and, equivalently,

$$\sum_{k=0}^{p-1}\left(1-\cos\left(\frac{2\pi i}{p}(r\cdot k)\right)\right)\cdot P^{[\nu]_p}(k)=0.$$

Since $1-\cos(\alpha)\geq 0$ for any α, and $P^{[\nu]_p}(k)>0$ for all k, we have $r\cdot k=0(\text{mod }p)$ for $k=0,1,2,\ldots,p-1$, which is possible only if $r=0(\text{mod }p)$.

Q.E.D.

Now, we can answer the question about convergence of probability distributions of residuals $\text{mod}(\nu^{(n)},p)$ as $n\to\infty$ for sums $\nu^{(n)}=\sum_{i=1}^{n}\nu_i$ $(n=1,2,\ldots)$ of independent random integers by the following statement.

Theorem 4.2

Let $\nu_1,\nu_2,\ldots,\nu_n,\ldots$ be a sequence of independent random integers (not necessarily equally distributed) such that for every prime $p\in\mathbb{P}$ the residuals $[\nu_i]_p$ $(i=1,2,\ldots)$ have probability distributions $P^{[\nu_i]_p}(k)>0$ for all $0\leq k\leq p-1$.

We assume that $\sup_{1\leq i\leq n,r\neq 0}\left|\Phi^{[\nu_i]_p}(r)\right|=M<1$ for $r\neq 0$. Then, the residuals of sums $\left[\nu^{(n)}\right]_p=\sum_{i=1}^{n}[\nu_i]_p$ are asymptotically uniformly distributed on G_p, for every $p\in\mathbb{P}$.

Proof.

We need to prove that $\lim_{n\to\infty} P^{\nu^{(n)}} = P^{\lambda}$, or simply that $\left[\nu^{(n)}\right]_p = \sum_{i=1}^{n}\left[\nu_i\right]_p \to \left[\lambda\right]_p$ (in probability) as $n \longrightarrow \infty$, where $[\lambda]_p$ is uniformly distributed on G_p .

We have $\Phi^{\nu^{(n)}} = \prod_{i=1}^{n}\Phi^{\nu_i}$ and $\left|\Phi^{\nu^{(n)}}(r)\right| = \prod_{i=1}^{n}\left|\Phi^{\nu_i}(r)\right| \le M^n \to 0$ as $n\to\infty$, for each $r \ne 0$.

This implies that $\lim_{n\to\infty}\Phi^{\left[\nu^{(n)}\right]_p}(r) = \Phi^{[\lambda]_p}(r) = \begin{cases} 1, \text{ if } r=0 \\ 0, \text{if } r\ne 0\end{cases}$, so that $\left[\nu^{(n)}\right]_p = \sum_{i=1}^{n}\left[\nu_i\right]_p \to \left[\lambda\right]_p$.

Thus, random variables $\left[\nu^{(n)}\right]_p$ are asymptotically uniformly distributed on $G_p = \mathbb{Z}_p$ as $n\to\infty$.

Q.E.D.

For a random variable $\nu\in\mathbb{N}$ we are interested in the vector of residuals $\vec{r}(\nu) = \left(r_1, r_2, \ldots r_{\pi(\nu)}\right)$, where $\pi(\nu)$ stands for number of primes $p \le \nu$.

Here $[\nu]_{p_i} = r_i = \mathrm{mod}(\nu, p_i)$ $\left(i = 1, 2, \ldots, \pi(\nu)\right)$ for all $p_i \le \nu$.

The asymptotic independence of residuals $[\nu]_{p_i} = r_i = \mathrm{mod}(\nu, p_i)$ $\left(i = 1, 2, \ldots, \pi(\nu)\right)$ is addressed in the following statement.

Theorem 4.4

All components of the vector of residuals $\vec{r}(\nu) = \left(r_1, r_2, \ldots r_{\pi(\nu)}\right)$are asymptotically independent random variables.

Proof.

Notice that the vector function $\mathrm{mod}\left(n, \vec{p}(\nu)\right) = \vec{r}(\nu) = \left(r_1, r_2, \ldots r_{\pi(\nu)}\right)$, where $\vec{p}(\nu) = \left(p_1, p_2, \ldots, p_{\pi(\nu)}\right)$, is periodic with a period $T(\nu) = \prod_{p\le\nu} p$ since $\mathrm{mod}(T(\nu), p) = 0$ for any $p \le \nu$. This implies that if x is a solution to the system of equations $\mathrm{mod}(x, p_i) = r_i$ $\left(1 \le i \le \pi(\nu)\right)$, then $y = x + T(\nu)$ is also a solution to the same system. We set $\nu = k(\nu)\cdot T(\nu) + r$, where $r = \mathrm{mod}\left(\nu, T(\nu)\right)$. Then,

$\mathrm{mod}(\nu, p_i) = \mathrm{mod}(r, p_i) = r_i$ and since the combination of residual values $\vec{r}(\nu) = (r_1, r_2, \ldots r_{\pi(\nu)})$ occurs $k(\nu)$ times in ν trials, then for the relative frequency $f(\nu, \vec{r}(\nu)) = \frac{k(\nu)}{\nu}$, we have: $\left|\frac{k(\nu)}{\nu} - \prod_{i \le \pi(\nu)} \frac{1}{p_i}\right| = \left|\frac{1}{T(\nu) + \frac{r}{k(\nu)}} - \frac{1}{T(\nu)}\right| \to 0$ as $\nu \to \infty$.

Q.E.D.

5. Distribution of Primes among Arithmetic Sequences

Denote $\Omega_m = \{C_{m,0}, C_{m,1}, \ldots, C_{m,m-1}\}$ a finite probability space with elementary events $\omega_r = C_{m,r}\ (r = 0,1,\ldots,m-1)$ that represent congruence classes modulo m. Then, we consider the corresponding algebra of events $\mathcal{F}_m$ generated by all subsets $\omega_r = C_{m,r}\ (r = 0,1,\ldots,m-1)$ of Ω_m and introduce a probability space $(\Omega_m, \mathcal{F}_m, P_m)$ where $P_m \in \mathcal{P}$ is a probability measure on Ω_m.

Since Ω_m provides a partition of $\mathbb{N}$, it induces a partition of the set of primes $\mathbb{P}$:

$$\mathbb{P} = \bigcup_{r=0}^{n-1} B_r, \quad B_r \cap B_t = \varnothing \quad (r \neq t) \text{ where } B_r = \mathbb{P} \cap C_{n,r}.$$

Notice, that for $b = k \cdot m + r \in C_{n,r}$ if numbers m and r have a common divider $d > 1$, then d divides b so that b cannot be prime. Therefore, we assume that m and r are co-prime numbers, that is $(m,r) = 1$. This is the case when $m = p \in \mathbb{P}$ is prime and $r = 1,2,\ldots,p-1$. Notice that $C_{p,r} = \{m \mid m = k \cdot p + r,\ k \in \mathbb{N}\}$ represents an arithmetic sequence with an initial term r and a common difference p. One of an old and known problems in Number Theory is the occurrence of prime numbers in arithmetic sequences $C_{p,r}$. L. Dirichlet proved in 1837 that $(p,r) = 1$ is the necessary and sufficient condition that there are infinitely many prime numbers in each of congruence classes $C_{p,r}\ (1 \leq r \leq p-1)$[2,5,10].

We illustrate in the tables below the occurrence of prime numbers within some arithmetic sequences, calculated with help of the author's scripts in R.

We denote the number of primes $p \leq x$ in class $C_{p,r}$ as $\pi_{p,r}(x)$. In examples below we provide the calculated values of primes occurred within finite arithmetic sequences: $m = p \cdot k + r \ (r = 1,2,\ldots,p-1;\ k = 1,2,\ldots,K)$ denoted as $\hat{\mathbb{P}}_{p,r}$ and the counts $\hat{\pi}_{p,r}$ of primes in these sequences.

Example 1. $K = 100,\ p = 5, r = 1$.

$$\hat{\mathbb{P}}_{5,1} = \left\{ \begin{matrix} 11 & 31 & 41 & 61 & 71 & 101 & 131 & 151 & 181 & 191 & 211 \\ 241 & 251 & 271 & 281 & 311 & 331 & 401 & 421 & 431 & 461 & 491 \end{matrix} \right\}$$

Example 2. $K = 100,\ p = 5, r = 2$

$$\hat{\mathbb{P}}_{5,2} = \left\{ \begin{matrix} 7 & 17 & 37 & 47 & 67 & 97 & 107 & 127 & 137 & 157 & 167 & 197 \\ 227 & 257 & 277 & 307 & 317 & 337 & 347 & 367 & 397 & 457 & 467 & 487\} \end{matrix} \right\}$$

Example 3. $K = 100,\ p = 5, r = 3$

$$\mathbb{P}_{5.3} = \left\{ \begin{matrix} 13 & 23 & 43 & 53 & 73 & 83 & 103 & 113 & 163 & 173 & 193 & 223 \\ 233 & 263 & 283 & 293 & 313 & 353 & 373 & 383 & 433 & 443 & 463 & 503 \end{matrix} \right\}$$

Example 4. $K = 100,\ p = 5, r = 4$

$$\hat{\mathbb{P}}_{5,4} = \left\{ \begin{matrix} 9 & 29 & 59 & 79 & 89 & 109 & 139 & 149 & 179 & 199 & 229 & 239 \\ 269 & 349 & 359 & 379 & 389 & 409 & 419 & 439 & 449 & 479 & 499 & \end{matrix} \right\}$$

Example 5. $K = 100,\ p = 13, r = 2$.

$$\hat{\mathbb{P}}_{13,2} = \left\{ \begin{matrix} 41\ 67\ 197\ 223\ 353\ 379\ 431\ 457\ 509 \\ 587\ 613\ 691\ 743\ 769\ \ 821\ 977\ 1237\ 1289 \end{matrix} \right\}$$

Here appears a problem: find a frequency distribution of prime numbers among congruence classes $\{C_{p,r} \mid r = 1,2,\ldots,p-1\}$. This problem had been solved in 1896 by La Valée-Puoussin, supporting the conjecture of even distribution of primes among the possible classes.

We illustrate the statement here (supported by computer calculations) that primes are asymptotically uniformly distributed over $\{C_{p,r} \mid r = 1,2,\ldots,p-1\}$, that is for each prime $p \geq 3$ and all remainder values $r = 1,2,\ \ldots\ ,p-1$, we should have

$$f_{p,r} = \frac{\hat{\pi}_{p,r}}{\sum_{t=1}^{p-1} \hat{\pi}_{p,t}} \approx \frac{1}{p-1} \tag{5.1}$$

This means that prime numbers asymptotically evenly populate congruence classes $\{C_{p,1}, C_{p,2}, \ldots, C_{p,p-1}\}$. See the Table 1 and Figure 1 below with the frequencies of prime numbers of the form $n \cdot p + r$ where $p = 31,\ n \leq 10^6,\ r = 1,2,\ldots,30$, in each of 30 congruence classes.

Table 5.1

$$\left[f_{p,r} \right] = \left\{ \begin{matrix} 0.03328864 & 0.03331891 & 0.03332309 & 0.03333561 & 0.03340451 & 0.03325263 \\ 0.03331108 & 0.03328133 & 0.03341547 & 0.03334918 & 0.03335649 & 0.03335858 \\ 0.03339302 & 0.03330638 & 0.03331265 & 0.03338363 & 0.03339929 & 0.03329647 \\ 0.03335545 & 0.03324845 & 0.03331839 & 0.03335075 & 0.03333352 & 0.03336223 \\ 0.03331734 & 0.03330221 & 0.03332935 & 0.03329803 & 0.03334240 & 0.03335492 \end{matrix} \right\}$$

Notice that $f_{31,r} \approx \frac{1}{30} = 0.0333\ldots$ for all $r\ (1 \leq r \leq 30)$.

We applied the Kolmogorov-Smirnov test to verify the claim about the uniform distribution of frequences of prime numbers among congruence classes $\left\{C_{p,1}, C_{p,2}, \ldots, C_{p,p-1}\right\}$ for $p = 31,\ n \le 10^6$:

```
ks.test(F6.31,runif)
One-sample Kolmogorov-Smirnov test
data:  F6.31
D = 0.86259, p-value = 1.221e-15
alternative hypothesis: two-sided
```

Consider now the question about divisibility of an unknown random number ν by a given integer m. We are looking for natural $m < n$ such that $n = k \cdot m$ for some $k \in \mathbb{N}$. This means that $n \in C_{m,0}$. Evaluation of probability that $\nu \in C_{m,0}$, where $\nu = n$, depends on the information about a given number n. There several assumptions are possible. We assume that given number n is a realization of a random variable ν distributed over $\mathbb{N}$ according to a probability distribution P. Then we evaluate the probability $P\left\{m \backslash \nu\right\} = P\left\{\nu \in C_{m,0}\right\}$ that natural m divides ν(that is $\operatorname{mod}(\nu, m) = 0$) . Generally, if $\operatorname{mod}(\nu, m) = r$, then $P\left\{\nu \in C_{m,r}\right\} = P\left(C_{m,r}\right),\ \sum_{r=0}^{m-1} P\left(C_{m,r}\right) = 1.$

Lemma 5.1

Let ν be a random variable that has a probability distribution on partition of $\mathbb{N}$ $\Omega_m = \left\{C_{m,0}, C_{m,1}, \ldots, C_{m,m-1}\right\}$ such that $P\left(\nu \in C_{m,0}\right) = P\left(C_{m,0}\right) = \dfrac{1}{m^s}$ ($s > 0$) for every $m \ge 2$.

Then events $C_{m_1,0}$ and $C_{m_2,0}$ are independent for any two co-prime numbers m_1 and m_2.

Proof.

We have $P\left(C_{m_1\cdot m_2,0}\right)=\dfrac{1}{\left(m_1\cdot m_2\right)^s}$. On the other hand, for any two co-prime number m_1 and m_2 we have $C_{m_1,0}\cap C_{m_2,0}=C_{m_1\cdot m_2,0}$, so that

$$P\left(C_{m_1,0}\cap C_{m_2,0}\right)=P\left(C_{m_1\cdot m_2,0}\right)=\frac{1}{m_1^s}\cdot\frac{1}{m_2^s}=P\left(C_{m_1,0}\right)\cdot P\left(C_{m_2,0}\right). \tag{5.2}$$

Q.E.D.

Remark 5.1

Notice that if a randomly chosen ν is evenly distributed within congruence classes $\Omega_m=\left\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\right\}$ then $P\left(\nu\in C_{m,0}\right)=\dfrac{1}{m}$ for every $m\geq 2$.

The assumption of the uniform distribution of ν on Ω_m within congruence classes corresponds to the minimum information (maximum uncertainty) about divisibility of an unknown random number ν by a given m (a maximum entropy value for a probability distribution on Ω_m for each natural $m\geq 2$).

The question is whether there exists a random variable $\nu:\Omega\to\mathbb{N}$ with its support on $\mathbb{N}$, uniformly distributed on each partition $\Omega_m=\left\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\right\}$ of $\mathbb{N}$ for every integer $m\geq 2$.

Here we demonstrate that that if ν has a Riemann Zeta probability distribution P_s , then $P\left\{\nu\in C_{m,0}\right\}=\dfrac{1}{\zeta(s)}\sum\limits_{k\in\mathbb{N}}\dfrac{1}{\left(k\cdot m\right)^s}=\dfrac{1}{m^s}\cdot\dfrac{1}{\zeta(s)}\cdot\sum\limits_{k\in\mathbb{N}}\dfrac{1}{k^s}=\dfrac{1}{m^s}$ for any $s>1$, and it satisfies the condition of the Lemma 2.1 above. This means independence of divisibility of Zeta distributed random number ν by any coprime numbers m_1 and m_2.

Remark 5.2

Meanwhile, obviously Zeta distribution $P_s\ (s>1)$ of ν on $\mathbb{N}$ does not provide a uniform distribution of ν on the partition $\Omega_m=\{C_{m,0},C_{m,1},\ldots,C_{m,m-1}\}$. Indeed, if ν has Zeta distribution, then $P_s\{\nu\in C_{m,0}\}=\frac{1}{m^s}$, and the uniform distribution

$P_s(C_{m,r})=\frac{1}{m^s}$ for all $r=0,1,2,\ldots,m-1$ would require $P_s(\mathbb{N})=\sum_{r=0}^{m-1}P_s(C_{m,r})=\frac{m}{m^s}=1$, which is possible only if $s=1$. But Zeta probability distribution P_s exists only for $s>1$.

For a prime $p\in\mathbb{P},\ p\geq 3$, the congruence classes $\{C_{p,r}\mid r=0,1,2,\ldots,p-1\}$ make a partition of $\mathbb{N}$. Obviously, $C_{p,0}=p\cdot\mathbb{N}$ and $\mathbb{P}_{p,0}=C_{p,0}\cap\mathbb{P}=\{p\}$. Therefore, $\{\mathbb{P}_{p,r}=C_{p,r}\cap\mathbb{P}\mid r=1,2,\ldots,p-1\}$ make a partition of the infinite set of primes $\mathbb{P}\setminus\{p\}$, and then, due to the 'pigeonhole principle', at least one of classes $C_{p,r}\ (1\leq r\leq p\text{-}1)$ must contain infinitely many prime numbers.

Notice, that for $\mathbb{P}$ the partition $\Omega_2=\{C_{2,0},C_{2,1}\}$ consists of two classes: the class of all even numbers $C_{2,0}$ and the class of all odd numbers $C_{2,1}$. Obviously, $C_{2,1}$ is the arithmetic sequence $\{n=2\cdot k+1\mid k\in\mathbb{N}\}$, which includes all prime numbers $p>2$.

Definition 5.1

A prime number $q\in\mathbb{P}, q\neq p$, we call $\mathbb{P}_{p,r}-prime$ if $q\in C_{p,r}$ that is if there exist $n\in\mathbb{N}$ and integer $r\ (1\leq r\leq p-1)$ such that $q=n\cdot p+r$.

In other words, set $\mathbb{P}_{p,r}$ consists of all $\mathbb{P}_{p,r}-\text{primes}$. Obviously, $\mathbb{P}_{p,r}$ is the intersection $C_{pr}\cap\mathbb{P}\setminus\{p\}$ of an arithmetic sequence $C_{pr}=\{n=k\cdot p+r\mid k\in\mathbb{N}\}$, $1\leq r\leq p-1$, and set of primes $\mathbb{P}\setminus p,\ p\in\mathbb{P}$.

The first goal is to prove that all sets of $\mathbb{P}_{p,r}$-prime numbers are not empty sets.

The second goal is to prove that every set $\mathbb{P}_{p,r}$ has infinitely many terms. The third goal is to prove that all sets $\mathbb{P}_{p,r}$ are evenly populated (in a statistical sense) by primes from the sets $\mathbb{P}\setminus\{p\}$. Due to the 'pigeonhole principle', at least for one value of r $(r=1,2,\ldots,p-1)$ set $\mathbb{P}_{p,r}\neq\varnothing$ contains infinitely many primes. As we have already mentioned above, the first two goals were completely achieved by L. Dirichlet [5, 2].

Denote $\pi_{p,r}(x)=\sum_{\mathbb{P}_{p,r}\cap[3,x]}1=\sum_{q\in\mathbb{P}_{p,r}}I_{[3,x]}(q)$ a number of $\mathbb{P}_{p,r}$-primes in the interval $[3,x]$.

Given a prime number q , the corresponding (reduced) vector of residuals $\vec{r}(q)=\left(r_1,r_2,\ldots,r_{\pi\left(\sqrt{q}\right)}\right)$ must have all non-zero components $r_i=\operatorname{mod}(q,p_i),\ 0\leq i\leq\pi\left(\sqrt{q}\right)$, $r_i=\operatorname{mod}(q,p_i),\ 0\leq i\leq\pi\left(\sqrt{q}\right)$, and obviously the complete vector of residuals $\vec{R}=\left(r_1,r_2,\ldots,r_{\pi(q)-1}\right)$ has also all non-zero components.

One of quite reasonable questions is how primes are distributed among arithmetic sequences. Since the sequence of all natural numbers is an example of an arithmetic sequence, the problem of occurrence of $\mathbb{P}_{p,r}$-primes within an arithmetic sequence $\{m=k\cdot p+r\mid k\in\mathbb{N}\}$ can be viewed as a generalization of the fundamental problem of distribution of primes in $\mathbb{N}$. The sequence of primes $\{p_i\}_{i\in\mathbb{N}}=\{2,3,5,\ldots\}$ is a deterministic sequence, and from probabilistic point of view we may consider $\mathbb{P}=\{p_i\}_{i\in\mathbb{N}}$ as a possible realization of a sequence $\{v_i(\omega)\}_{i\in\mathbb{N}}$ of random variables $v_i(\omega)=p_i$ defined on an appropriate probability space $(P,\mathcal{F},\Omega)$. Similarly, we may consider a sequence $\mathbb{P}_{p,r}=\{q_1,q_2,\ldots,q_n,\ldots\}$, where $q_k=n_k\cdot p+r,\ (n_1<n_2<\cdots<n_k<\cdots)$ and $q_k\in\mathbb{P}$, so that $v_i(\omega)=q_i\in\mathbb{P}_{p,r}$.

Then, we can evaluate probability $P\left\{v_i(\omega)=q_i\in\mathbb{P}_{p,r}\right\}$ by the relative frequency:

$$P\left\{v \text{ is a } \mathbb{P}_{p,r}\text{-prime}|\ v\in\mathbb{P}\cap\left[p+r,x\right]\right\}\approx\frac{\pi_{p,r}(x)}{\sum_{t=1}^{p-1}\pi_{p,r}(x)} \tag{5.3}$$

Denote $\xi_{p,r}(m)=\begin{cases}1 \text{ if } m \text{ is } \mathbb{P}_{p,r}\text{-prime}\\ 0, \text{ otherwise}\end{cases}$. Then, $\hat{\pi}_{p,r}(x)=\sum_{v\le x}\xi_{p,r}(v)$,

where $\xi_{p,r}(v)=\begin{cases}1 \text{ if } \left(v>p, v\in\mathbb{P}, \operatorname{mod}(v,p)=r\right)\\ \xi_{p,r}(v)=0 \text{ otherwise}\end{cases}$.

Let $\left\{v_n\right\}_{n\ge2}$ be a sequence of random variables $v_n=n\cdot\xi(n)$ where $\xi(n)=\begin{cases}1 \text{ if } n\in\mathbb{P}\\ 0 \text{ otherwise}\end{cases}$

The following theorem answers the question about the asymptotic behavior of residual values $\operatorname{mod}(v_n,p)=\left[v_n\right]_p$. It turns out that asymptotically (as $n\to\infty$) all values of residuals $\left[n\right]_p=r\ (0\le r\le p-1)$ are equally likely to occur.

Theorem 5.1

Let $\left(\xi(n)\,|\,n=1,2,\ldots\right)$ be a sequence of prime number indicators $\xi(n)=\begin{cases}1 \text{ if } n\in\mathbb{P}\\ 0 \text{ otherwise}\end{cases}$ and $\mathbb{P}_{pr}=C_{pr}\cap\mathbb{P}\setminus\left\{p\right\}$ a set of primes , except for p , in the arithmetic sequence $C_{pr}\ \left(1\le r\le p-1\right)$. For $\xi_{p,r}(n)=\begin{cases}1 \text{ if } n \text{ is } \mathbb{P}_{p,r}\text{-prime}\\ 0, \text{ otherwise}\end{cases}$ and $v_n=n\cdot\xi_{pr}(n)$, denote $P^{\left[v_n\right]_p}(r)=P\left\{\left[v_n\right]_p=r\right\}$ the probability distribution of $\left[v_n\right]_p$ on $\left\{C_{p,r}\,|\,r=1,2,\ldots,p-1\right\}$. Then, $\left[v_n\right]_p\to v_0$ in probability, where random variable v_0 has a uniform distribution on $\left\{01,2,\ldots,p-1\right\}$.

Proof.

Consider a characteristic function $\Phi^{[\nu_n]_p}(r) = Ee^{ir\cdot[\nu_n]_p} = \sum_{t=0}^{p-1} e^{ir\cdot t} P^{[\nu_n]_p}(t)$ of probability distribution $P^{[\nu_n]_p}(r) = P\left\{\left[\nu_n\right]_p = r\right\} = P\left\{\xi_{pr}(\nu_n) = 1\right\}$ of $\left[\nu_n\right]_p$ on $\left\{C_{p,r} \mid r = 1,2,\ldots,p-1\right\}$.

Due to the Cramer's model approximation for probabilities of primes,

$P\left\{\xi_{p,r}(\nu_n) = 1\right\} = P\left\{C_{p,r}(\nu_n) \cap \left(\nu_n \in \mathbb{P}\right)\right\} \le P\left\{\nu_n \in \mathbb{P}\right\} \approx \frac{1}{\ln n}\ (n \ge 2)$ for all $r\ \left(1 \le r \le p-1\right)$

We have then $\left|\Phi^{[\nu_n]_p}(r)\right| \le \frac{p}{\ln n}$. This implies: $\lim_{n\to\infty} \Phi^{[\nu_n]}(r) = \begin{cases} 1 \text{ if } r = 0 \\ 0 \text{ otherwise} \end{cases} = \Phi^{\nu_0}(r)$

where Φ^{ν_0} is a characteristic function of ν_0 uniformly distributed on $\left\{1,2,\ldots,p-1\right\}$.

Q.E.D.

Theorem 5.2

By using the Cramer's model approximation,

$P\left\{\xi_{p,r}(\nu) = 1 \mid \nu = n\cdot p + r \in \mathbb{P}_{p,r}, \nu < N\right\} = \frac{1}{\ln(n\cdot p + r)} = \lambda_{p,r}(n)$. Then,

$$E\left\{\hat{\pi}_{p,r}(x)\right\} = \sum_{n\le x} \lambda_{p,r}(n) \sim \int_2^x \lambda_{p,r}(t)dt = \int_2^x \frac{dt}{\ln(t\cdot p + r)} = \frac{1}{p}\cdot\left[Li(x\cdot p + r) - Li(p + r)\right] \qquad (5.4))$$

Assuming that terms in the sequence $\left(\xi(n) \mid n = 1,2,\ldots\right)$ are uncorrelated, we have

$$Var\left\{\hat{\pi}_{p,r}(x)\right\} = \sum_{n\le x} \lambda_{p,r}(n)\cdot\left(1 - \lambda_{p,r}(n)\right) \sim \int_3^x \lambda_{p,r}(t)\cdot\left(1 - \lambda_{p,r}(t)\right)dt \qquad (5.5)$$

Proof

Due to the Cramer's model approximation for probabilities of primes:

$$P\left\{\xi_{p,r}(\nu) = 1 \mid \nu = n\cdot p + r \in \mathbb{P}_{p,r}, \nu < N\right\} = \frac{1}{\ln(n\cdot p + r)} = \lambda_{p,r}(n) \qquad (5.6)$$

Thus, mathematical expectation and variance of $\xi_{p,r}(\nu)$ given $\nu = n\cdot p + r$ can be approximated as

$$E\left\{\xi_{p,r}(\nu)\} \mid \nu = n\cdot p + r\right\} = \lambda_{p,r}(n),\ Var\left\{\xi_{p,r}(\nu)\} \mid \nu = n\cdot p + r\right\} = \lambda_{p,r}(n)\cdot\left(1 - \lambda_{p,r}(n)\right). \qquad (5.7)$$

This implies:

$$
\begin{aligned}
E\left\{\hat{\pi}_{p,r}(x)\right\} &= \sum_{\nu\le x} E\left\{\xi_{p,r}(\nu) \mid \nu = n\cdot p + r\right\} = \sum_{p+r\le n\le x} \lambda_{p,r}(n) \\
Var\left\{\hat{\pi}_{p,r}(x)\right\} &= \sum_{\nu\le x} Var\left\{\xi_{p,r}(\nu) \mid \nu = n\cdot p + r\right\} = \sum_{p+r\le n\le x} \lambda_{p,r}(n)\cdot\left(1-\lambda_{p,r}(n)\right)
\end{aligned} \tag{5.8}
$$

Using (30), we can approximate the mathematical expectation and variance of $\xi_d(\nu)$ in the integral form by using the Eulerian integral approximation.

We have then, $E\hat{\pi}(x) = \sum_{2\le n\le x} \frac{1}{\ln n} \sim \int_2^x \frac{dt}{\ln t} = Li(x) \approx \pi(x)$.

Notice by the way that $\pi_{2,1}(x) = \pi(x)$.

For the general case of arithmetic sequences, we can write:

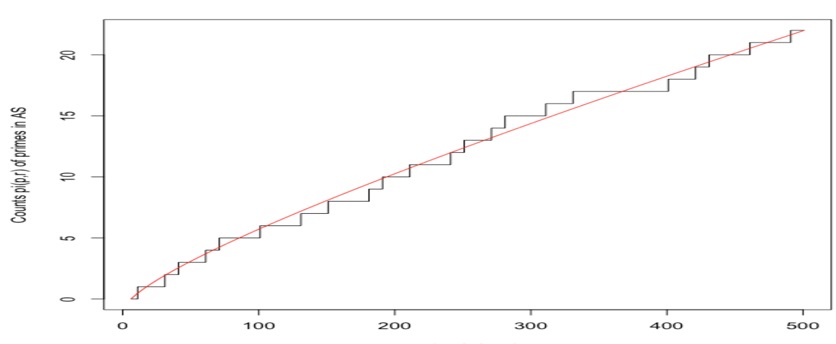
$$
\begin{aligned}
E\left\{\hat{\pi}_{p,r}(x)\right\} &= \sum_{n\le x} \lambda_{p,r}(n) \sim \int_2^x \lambda_{p,r}(t)dt = \int_2^x \frac{dt}{\ln(t\cdot p + r)} = \frac{1}{p}\cdot\left[Li(x\cdot p + r) - Li(p+r)\right] \\
Var\left\{\hat{\pi}_{p,r}(x)\right\} &= \sum_{n\le x} \lambda_{p,r}(n)\cdot\left(1-\lambda_{p,r}(n)\right) \sim \int_3^x \lambda_{p,r}(t)\cdot\left(1-\lambda_{p,r}(t)\right)dt
\end{aligned} \tag{5.9}
$$

Q.E.D.

The figure below illustrates distribution of primes $\hat{\pi}_{p,r}(x)$ and its mathematical expectation $E\left\{\hat{\pi}_{pr}(x)\right\}$ in the arithmetic sequence $\{p\cdot k + r\}$ for the given values of p, r and k.

Graphs of $\hat{\pi}_{pr}(x)$ (step-function) and $E\left\{\hat{\pi}_{pr}(x)\right\}$ (solid line)

for $p = 5, r = 1, k = 1,2,\ldots,100$.



6. Additive Walks and Distribution of Twin- and d- Primes

Consider an additive rule to generate stochastic or deterministic sequences of positive integers:

$$\nu(1) = 0,\ \nu(k+1) = \nu(k) + \xi(k+1), \tag{6.1}$$

where $\xi(k) = \begin{cases} k & \text{if } k \in \mathbb{P} \\ 0 & \text{otherwise} \end{cases}$ for all $k = 1,2,3\ldots$

This approach leads to 'additive models' of random walks in the study of prime numbers distribution. Though the sequence $\nu(k)\ (k = 1,2,\ldots,)$ generated recursively is deterministic, each step of the 'walk' (6.1) can result either in a prime $\nu(k) = p_k \in \mathbb{P}$, or in 0 (if k is a composite number), and the differences ('gaps') $\xi(k) = \Delta p_k = p_{k+1} - p_k$ between consecutive primes look very sporadic and hard to predict. It is well known that the gaps between two consecutive primes $p_k \geq 3$ and p_{k+1} can be as small as 2 (for twin primes) or arbitrary big (see the table below). Indeed, in the sequence of $n-1$ consecutive integers $\{n!\ +\ k \mid 2 \leq k \leq n\}$ each integer $n!\ +\ k$ is divisible by k, and therefore this sequence does not include primes. This means that there are consecutive prime numbers p_i and p_{i+1} such that $p_i < n!\ +\ 2$ and $p_{i+1} > n!\ +\ n$, which implies that $\Delta p_i = p_{i+1} - p_i \geq n$. The next definition is a generalization of the notion of twin primes.

Definition 6.1

We call prime numbers $p < p'$ *consecutive* if there is no prime q between them (that is no prime q such that $p < q < p'$). A prime number p we call d*-prime* if p, p' are consecutive primes and $p' = p + d$.

Notice that the number $d = p_{i+1} - p_i = g_i$ for a d-prime p_i is called a "gap between two successive primes" (see the article "Prime gap" in [8]).

For example, p is a 2-prime, if and only if p and $p+2$ are twin primes, since for $p \geq 3$ twin primes p and $p+2$ are automatically consecutive. Let us denote $D\mathbb{P}_d = \{p \mid p \text{ and } p+d \text{ are consecutive primes}\}$, the set of d-primes (that is prime numbers p such that p and $p+d$ are consecutive primes).

For example, $D\mathbb{P}_1 = \{2\}$; the set of twin primes is $D\mathbb{P}_2 = \{3,5,11,17,29,41,...\}$.

One of famous conjectures is that the set $D\mathbb{P}_2$ is infinite.

Table 6.1. *d-primes for $d = 2,4,6$, among all primes $p < 200$*

$D\mathbb{P}_2$	3 5 11 17 29 41 59 71 101 107 137 149 179 191 197
$D\mathbb{P}_4$	7 13 19 37 43 67 79 97 103 109 127 163 193
$D\mathbb{P}_6$	23 31 47 53 61 73 83 131 151 157 167 173

Notice that $D\mathbb{P}_d = \varnothing$ for all odd $d > 1$ and *the first conjecture is that* $D\mathbb{P}_d \neq \varnothing$ *for all even values of* $d \geq 2$. Obviously, $D\mathbb{P}_d \cap D\mathbb{P}_{d'} = \varnothing$ for all $d \neq d'$ and

$D\mathbb{P}_1 \cup \left[\bigcup_{even\ d=2}^{\infty} D\mathbb{P}_d \right] = \mathbb{P}$, that is $\{D\mathbb{P}_d\}_{d \in 2\mathbb{N}}$ makes a partition of the set of primes.

This means that any prime number p is a d-prime for an appropriate d. Indeed, due to the Euclid theorem, there are infinitely many prime numbers, therefore for any prime p there exist the next (that is consecutive) prime p' and $p \in D\mathbb{P}_d$ where $d = p' - p$.

The second conjectire is that every $D\mathbb{P}_d$ *is an infinte set for all even values of* $d \geq 2$.

Lemma 6.1

For a positive even integer d and a prime number p such that $p > d$, we have $p \in D\mathbb{P}_d$ if and only if $p \in C_{d,r}$ and $(p+d) \in C_{d,r}$, where r is an odd number such that $1 \le r \le d-1$ and $p = k \cdot d + r,\ p + d = (k+1) \cdot d + r$.

Proof.

Let p_i and p_{i+1} be two consecutive prime numbers. We have

$p_i = k_i \cdot d + r_i,\ p_{i+1} = k_{i+1} \cdot d + r_{i+1}$,

where $1 \le r_i \le d-1,\ 1 \le r_{i+1} \le d-1$. Then, $p_{i+1} = p_i + d$ implies

$\Delta\, p_i = (k_{i+1} - k_i) \cdot d + (r_{i+1} - r_i) = d$.

Since $|r_{i+1} - r_i| < d$, we should have $r_{i+1} = r_i = r$ and $k_{i+1} - k_i = 1$.

Thus, $p_i = k_i \cdot d + r,\ p_{i+1} = (k_i + 1) \cdot d + r$ where r is an odd number and $r \ge 1$.

Q.E.D.

Remark 6.1

Since for each even number d the finite number of congruence classes $\left\{C_{d,r} \cap \mathbb{P} \,|\, 1 < \mathit{even}\ r < d-1\right\}$ make a partition of the infinite set of all primes $\mathbb{P}$, at least one of classes $C_{d,r}$ must contain infinitely many prime numbers.

Prime numbers populate the sets $D\mathbb{P}_{d,N} = D\mathbb{P}_d \cap [2, N]$ not evenly for different even integers d, as illustrated by the histogram below for $N = 10^9$. Computer calculations show so far that the most frequent value of consecutive prime gaps is $d = 6$.

According to the Prime Number Theorem [10, p.133], the counting function of primes on $\mathbb{N}$ is given by the asymptotic formula

$$\pi(x) = \sum_{p \in \mathbb{P} \cap [2,x]} 1 \sim Li(x) = \int_2^x \frac{dt}{\ln t} \tag{6.2}$$

This leads to the heuristic assumption about the probability

$$P\{p \in [x-1, x]\} \sim \int_{x-1}^{x} \frac{dt}{\ln t} \sim \frac{1}{\ln x}.$$

According to the Cramér's model, occurrences of primes in $\mathbb{N}$ are controlled by the sequence of *independent Bernoulli variables* $\{\xi_n\}_{n\geq 3}$ such that $\xi_n = \begin{cases} n & \text{if } n \in \mathbb{P} \\ 0 & \text{otherwise} \end{cases}$

and

$$P\{\xi_n = 1\} = \frac{1}{\ln n},\ P\{\xi_n = 0\} = 1 - \frac{1}{\ln n} \ \text{ for all } \ n \in \mathbb{N} \cap \{n \geq 3\}. \tag{6.3}$$

As we know, the sequence of primes $\{p_n\}_{n\in\mathbb{N}}$ is deterministic and is recursively controlled by the corresponding vectors of residual

$$\vec{r}(n) = \left(r_1, r_2, \ldots, r_{\pi\left(\sqrt{n}\right)}\right), \text{ where } r_i = \text{mod}(n, p_i),\ i = 1, 2, \ldots, \pi\left(\sqrt{n}\right).$$

Therefore, in contrast to the Cramér's model, the terms of a more adequate sequence of random variables $\{\xi_n\}_{n\geq 3}$ cannot be independent, since ξ_n must be equal to 0 for all even n. Indeed, any prime $p > 2$ is an odd number, and all primes, except for $p_1 = 2$, belong to the set of odd numbers (that is to the congruence class $C_{2,1}$ of $\mathbb{N}$). Denote $\pi_d(x) = \sum_{D\mathbb{P}_d \cap [2,x]} 1 = \sum_{p \in D\mathbb{P}_d} I_{[2,x]}(p)$ a number of d-primes in the interval $[2,x]$. Given a prime number p, the corresponding vector of residuals $\vec{r}(p) = \left(r_1, r_2, \ldots, r_{\pi\left(\sqrt{p}\right)}\right)$ must have all non-zero components $r_i = \text{mod}(p, p_i),\ 0 \leq i \leq \pi\left(\sqrt{p}\right)$ and obviously the complete vector of residuals $\vec{R} = \left(r_1, r_2, \ldots, r_{\pi(p)-1}\right)$ has also all non-zero components.

One of quite reasonable questions is how frequently d-primes may occur among all prime numbers? We can evaluate the empirical probability of d-primes by the relative frequency:

$$P\left\{\nu \text{ is a } d\text{-prime}|\ \nu \in \mathbb{P} \cap \left[2,x\right]\right\} \approx \frac{\pi_d(x)}{\pi(x)} \tag{6.4}$$

Denote $\xi_d(n) = \begin{cases} 1 \text{ if } n \text{ is } d\text{-prime} \\ 0, \text{ otherwise} \end{cases}$. Then $\pi_d(x) = \sum_{n\le x} \xi_d(n)$,

where $\xi_d(\nu) = 1$ if $\nu = p_i$ and $\nu + d = p_{i+1}$ are consecutive prime numbers. Thus,

$$\pi_d(\nu + d) = \pi_d(\nu) + \xi_d(\nu) \quad (i = 1,2,\ldots). \tag{6.5}$$

Assuming the Cramer's assumption of independence of consecutive primes, we have:

$$\begin{aligned}
&P\left\{\xi_d(\nu) = 1\right\} = P\left\{\nu \text{ and } \nu + d \text{ are consecutive primes}\right\} \\
&= P\left\{\nu \text{ and } \nu + d \text{ are prime numbers with no primes in the open interval } \left(\nu, \nu + d\right)\right\} \\
&= P\left\{\nu \in \mathbb{P}\right\} \cdot P\left\{\bigcap_{i=1}^{d-1}\left\{(\nu + i) \notin \mathbb{P}\right\}\right\} \cdot P\left\{(\nu + d) \in \mathbb{P}\right\}
\end{aligned}$$

Then, $P\left\{\bigcap_{i=1}^{d-1}\left\{(\nu + i) \notin \mathbb{P}\right\}\right\} = \prod_{i=1}^{d-1}\left(1 - P\left\{(\nu + i) \in \mathbb{P}\right\}\right)$. Following the Cramér's model

assumption: $P\left\{\nu \in \mathbb{P} \mid \nu = n\right\} = \dfrac{1}{\ln n}$, we obtain

$$P\left\{\xi_d(\nu) = 1 \mid \nu = n\right\} = \frac{1}{\left(\ln n\right) \cdot \left(\ln(n + d)\right)} \cdot \prod_{i=1}^{d-1}\left(1 - \frac{1}{\ln(n + i)}\right) = \Psi(n,d) \tag{6.6}$$

Denoting $\phi(n,d) = \prod_{i=1}^{d-1}\left(1 - \dfrac{1}{\ln(d + i)}\right)$, we write the function $\Psi(n,d)$ in (6.6) as

$$\Psi(n,d) = \frac{\phi(n,d)}{\ln(n) \cdot \ln(n + d)} \tag{6.7}$$

Thus, mathematical expectation and variance of $\xi_d(\nu)$ given $\nu = n$ can be approximated as

$E\left\{\xi_d(\nu)\}\,|\,\nu=n\right\}=\Psi(n,d),\;\; Var\left\{\xi_d(\nu)\}\,|\,\nu=n\right\}=\Psi(n,d)\cdot\left(1-\Psi(n,d)\right)$. This implies:

$$\begin{aligned} E\left\{\pi_d(x)\right\} &= \sum_{n\le x} E\left\{\xi_d(\nu)\}\,|\,\nu=n\right\} = \sum_{n\le x}\Psi(n,d) \\ Var\left\{\pi_d(x)\right\} &= \sum_{n\le x} Var\left\{\xi_d(\nu)\}\,|\,\nu=n\right\} = \sum_{n\le x}\Psi(n,d)\cdot\left(1-\Psi(n,d)\right) \end{aligned} \tag{6.8}$$

Using (6.6 and 6.7), we can approximate the mathematical expectation and variance of $\xi_d(\nu)$ in the integral form:

$$\begin{aligned} E\left\{\pi_d(x)\right\} &= \sum_{n\le x}\Psi(n,d) \sim \int_2^x \frac{\phi(t,d)}{\ln(t)\cdot\ln(t+d)}dt \\ Var\left\{\pi_d(x)\right\} &\sim \int_2^x \frac{\phi(t,d)}{\ln(t)\cdot\ln(t+d)}\cdot\left(1-\frac{\phi(t,d)}{\ln(t)\cdot\ln(t+d)}\right)dt \end{aligned} \tag{6.9}$$

The comparison of $\pi_d(x)$ distribution with its mathematical expectation $E\pi_d(x)$ is given in the tables below computed for $d=2,4,6,8,10,12$ and x changing in steps: $10^1, 10^2, \ldots, 10^8$.

Number $\pi_d(x)$ of d-primes for $p\le x$

d:	2	4	6	8	10	12	14	16
x:								
10^2	8	7	7	1	0	0	0	0
10^3	35	40	44	15	16	7	7	0
10^4	205	202	299	101	119	105	54	33
10^5	1224	1215	1940	773	916	964	484	339
10^6	8169	8143	13549	5569	7079	8005	4233	2881
10^7	58980	58621	99987	42352	54431	65513	35394	25099
10^8	440312	440257	768752	334180	430016	538382	293201	215804

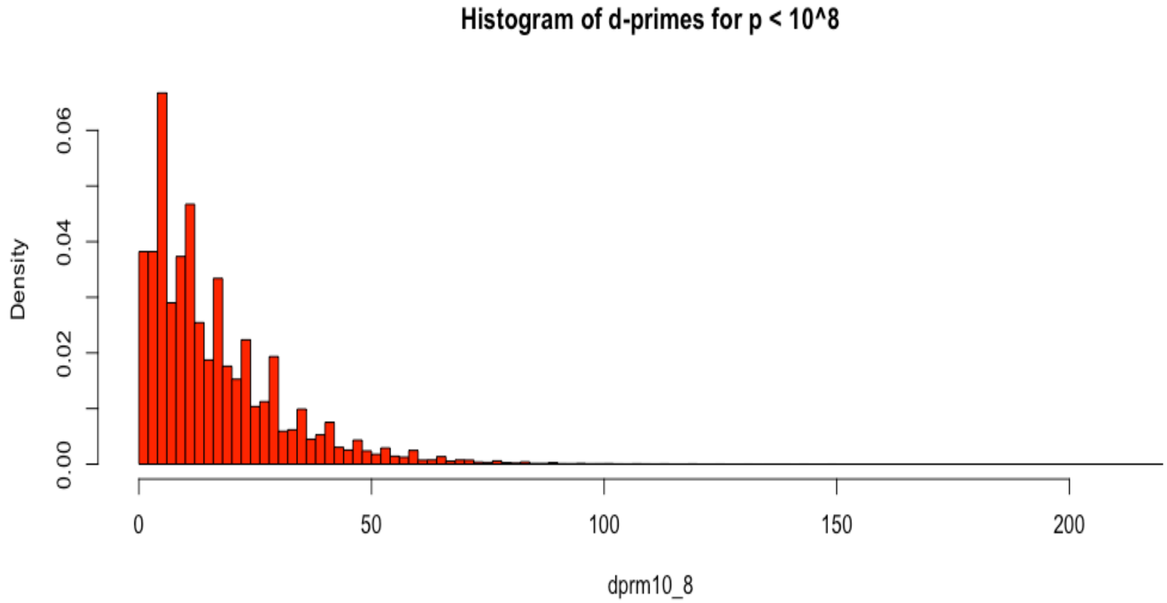
Expectation $E\pi_d(x)$ of numbers of d-primes for $p \le x$

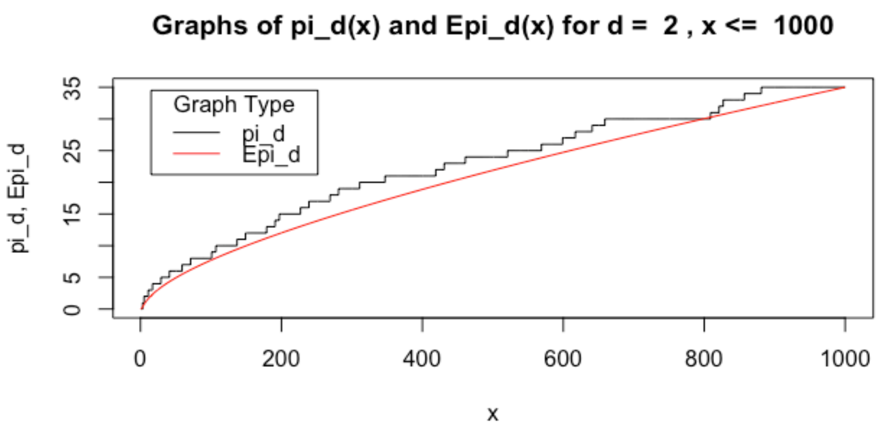
x: d:	2	4	6	8	10	12	14	16
10^2	5	5	5	1	0	0	0	0
10^3	27	32	36	13	13	6	6	0
10^4	177	175	261	88	104	92	47	29
10^5	1100	1093	1748	697	827	871	437	307
10^6	7510	7487	12464	5124	6515	7371	3898	2653
10^7	55001	54667	93255	39505	50776	61125	33026	23422
10^8	414685	414638	724062	314770	405047	507165	276210	203311

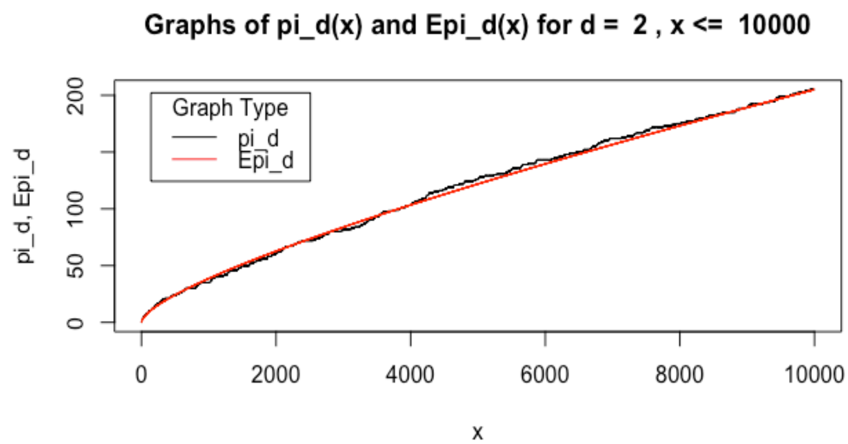
The quality of prediction of $\pi_d(x)$ by its expectation $E\pi_d(x)$ is given by the measure of relative error $R_d(x) = \frac{\pi_d(x) - E\pi_d(x)}{E\pi_d(x)}$ as illustrated by the table below.

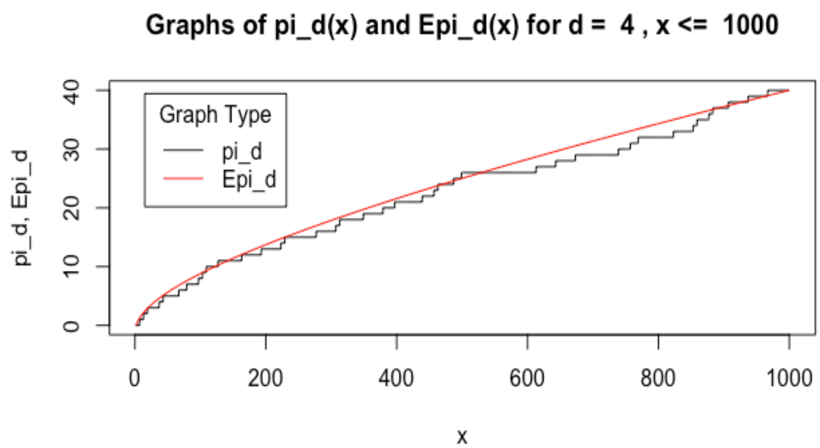
Relative errors $\varepsilon_d(x) = \frac{\pi_d(x) - E\pi_d(x)}{E\pi_d(x)}$

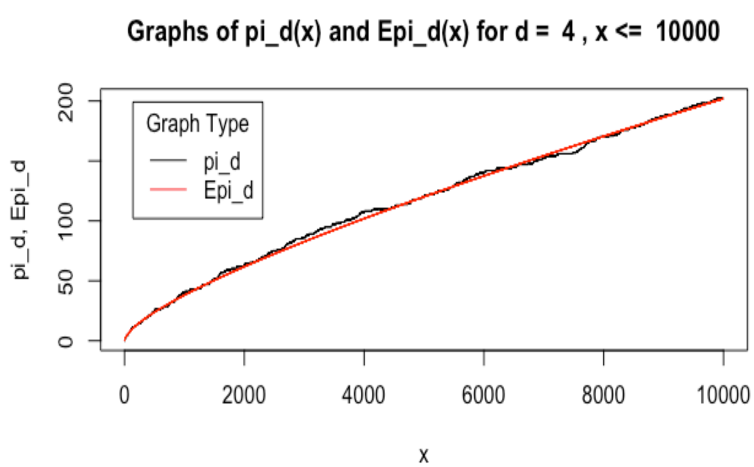
d: x:	2	4	6	8	10	12	14	16
10^2	0.628	0.445	0.344	0.284	NaN	NaN	NaN	NaN
10^3	0.284	0.238	0.219	0.191	0.196	0.198	0.196	NaN
10^4	0.160	0.155	0.147	0.144	0.144	0.138	0.141	0.134
10^5	0.112	0.112	0.110	0.108	0.108	0.107	0.107	0.106
10^6	0.088	0.088	0.087	0.087	0.086	0.086	0.086	0.086
10^7	0.072	0.072	0.072	0.072	0.072	0.072	0.072	0.072
10^8	0.062	0.062	0.062	0.062	0.062	0.062	0.062	0.061


Histogram of d-primes for p < 10^8
Density
0.00
0.02
0.04
0.06
0
50
100
150
200
dprm10_8


Graphs of pi_d(x) and Epi_d(x) for d = 2 , x <= 1000
pi_d, Epi_d
0
5
15
25
35
Graph Type
pi_d
Epi_d
0
200
400
600
800
1000
x


Graphs of pi_d(x) and Epi_d(x) for d = 2 , x <= 10000
Graph Type
pi_d
Epi_d
pi_d, Epi_d
x


Graphs of pi_d(x) and Epi_d(x) for d = 4 , x <= 1000
Graph Type
pi_d
Epi_d
pi_d, Epi_d
x


Graphs of pi_d(x) and Epi_d(x) for d = 4 , x <= 10000
Graph Type
pi_d
Epi_d
pi_d, Epi_d
x

Theorem 6.1

For each even value of $d \geq 2$ there are infinitely many consecutive prime numbers with a gap equal to d (so that every $D\mathbb{P}_d$ is an infinte set for all even values of $d \geq 2$).

Proof.

This statement could be proved by using the equivalence:

$$E\{\pi_d(x)\} = \sum_{n \leq x} \Psi(n,d) \sim \int_2^x \frac{\phi(t,d)}{\ln(t) \cdot \ln(t+d)} dt \text{ as } x \to \infty.$$

Indeed, if we assume that there exists $x_{\max}$ such that $\pi_d(x) = \pi_d(x_{\max})$ for all $x \geq x_{\max}$, then $\pi_d(x)$ becomes constant for sufficiently large values of x. But this contradicts the above equivalence since function $F(x) = \int_2^x \frac{\phi(t,d)}{\ln(t) \cdot \ln(t+d)} dt$ is strictly increasing for all $x > 2$ because its derivative $F'(x) = \frac{\phi(x,d)}{\ln(x) \cdot \ln(x+d)} > 0$ for all even $d \geq 2$ and $x > 2$.

Q.E.D.

7. An additive model for the Goldbach Strong Conjecture

According to the conjecture stated by Goldbach in his letter to Euler in 1742, "every even number $2m \geq 6$ is the sum of two odd primes" [1]. Regardless numerous attempts to prove the statement, supported in our days by computer calculations up to 4×10^{18} , it remains unproven till now. In this notice we try to 'solve' the puzzle in the framework of Probability Theory, by using the H. Cramér's assumption of independence of primes occurrence in the sequence of natural numbers $\mathbb{N}$.

We consider the so-called *Goldbach function* $G(2m)$ that denotes the number of presentations of an even integer in the form: $2m = p + p'$ where p, p' are prime numbers (called G-primes).

A choice of a G-prime p for every $m \geq 3$ is considered as a realization of $G(2m,\nu)$ for a random variable ν with Zeta probability distribution.

We have then, $2m = \nu + \nu'$ where $\nu \in \mathbb{P}$, $\nu' \in \mathbb{P}$. The calculations for the available range of ν values show that the number of representations $G(2m)$ of an even integer in the form $2m = p + p'$ where p, p' are primes, increases when m increases and becomes larger for the larger values of m.

Definition 7.1

A prime number $p \in \mathbb{P}$, $p \leq m$, we call a G_m- *prime* if there exist an even number $2m \geq 6$ and a prime number $p' \in \mathbb{P}$ such that $2m = p + p'$.

The set of all G_m-*primes* for a given m we denote $G_m\mathbb{P}$.

The Goldbach Conjecture (GC) asserts that all sets $G_m\mathbb{P}$ are not empty:

$G_m\mathbb{P} \neq \varnothing$ for all $m \geq 3$, and it can be stated as $G(2m) = |G_m\mathbb{P}| > 1$ for all $m \geq 3$, where $|A|$ denotes a number of elements in a finite set A.

For all $m \geq 3$ we focuse on the number $G(2m)$ of *Goldbach* m*-primes,* or G_m*-primes*, which are primes p such that a difference $2m - p$ is again a prime number. This means that. $G_m\mathbb{P} = \{p_k \mid 1 \leq k \leq G(2m), p_k \in \mathbb{P}, 2m - p_k \in \mathbb{P}\}$.

If there exists $m \geq 3$ such that $G_m\mathbb{P} = \varnothing$, then $G(2m) = 0$. In the context of the *Goldbach conjecture* we are interested in evaluation of probability distribution for the numbers $G(2m)$ of primes in the representations for an even number $2m$ in the form

$$2m = p + p', \text{ where } p, p' \in \mathbb{P} . \tag{7.1}$$

To incorporate a probabilistic approach in this context we consider a choice of a G_m- prime for every $m \geq 3$ as realization $\nu_k = p \in \mathbb{P}$ of one of random variables ν_k, $3 \leq k \leq m$, with a certain probability distributions $P_m \in \mathcal{P}$, where $\mathcal{P}$ is a set of all σ-additive probability measures on $\mathbb{N}$.

We have then for each k, $3 \leq k \leq m$, $2m = \nu_k + \nu_k'$ where $\nu_k \in \mathbb{P}$, $\nu_k' \in \mathbb{P}$.

There appear some questions related to the 'modelling' of primes by sequence of random variables $(\nu_k)_{k\in\mathbb{N}}$.

1. How well a chosen probability distribution of terms in a sequence $(\nu_k)_{k\in\mathbb{N}}$ represents the natural sequence $(p_k)_{k\in\mathbb{N}}$ of the set of prime numbers $\mathbb{P}$.
2. What kind on conditions (restrictions) we need to impose on $(\nu_k)_{k\in\mathbb{N}}$ for an adequate 'modelling' of the sequence $(p_k)_{k\in\mathbb{N}}$ of primes.
3. As we have demonstrated above in Chapter 3 (Lemma 3.2), the sequence $(\xi_k)_{k\in\mathbb{N}}$ in the Cramér's model is *asymptotically pairwise Bernoullian,* so that

$$\left| V\{X_n\} - \frac{1}{n^2}\sum_{k=1}^{n} V\{\xi_k\} \right| = O\left(\frac{1}{\ln N}\right) \quad \text{as} \quad D_n = O\left(\frac{1}{\ln N}\right) \text{ for all } n > N.$$

4. Would the proved conclusion regarding the sequence of random variables $(\nu_k)_{k\in\mathbb{N}}$ be valid for $\mathbb{P}$.

Denote vector $\nu^{(m)} = (\nu_k \mid 3 \le k \le m)$ and consider $G(2m)$ as a realization of $G(2m,\nu^{(m)})$. Defining

$$\gamma_m(n) = \begin{cases} 1 & \text{if } n \in G_m\mathbb{P} \\ 0, & \text{otherwise} \end{cases} \text{ with } P\{\gamma_m(\nu_k) = 1 \mid \nu_k = n\} = P_m\{\nu_k \in G_m\mathbb{P} \mid \nu_k = n\}, \tag{7.2}$$

we have

$$G(2m,\nu^m) = \sum_{3\le n\le m} \gamma_m(n) = \sum_{k=3}^{m} \gamma_m(\nu_k) \tag{7.3}$$

Assuming $(\nu_k)_{k\in\mathbb{N}}$ to be a sequence of independent random variables on probability space $(\Omega,\mathcal{F},P)$ with $P\{\xi(n)=1\} = P\{\nu_k = n \in \mathbb{P}\}$, we have independence of terms in the sequence $(\xi(n) \mid n \in \mathbb{N})$, and therefore

$$\begin{aligned} &P\{\gamma_m(\nu_m) = 1 \mid \nu_m = n\} = P\{(\nu_m \in \mathbb{P}) \cap ((2m-\nu_m) \in \mathbb{P}) \mid \nu_m = n\} \\ &= P\{\xi(n)=1\} \cdot P\{\xi(2m-n)=1\} \end{aligned} \tag{7.4}$$

Validity of the choice of probabilities in the Cramer's model is supported by formula (2.2) in Theorem 2.3, and by Merten's 1[st] and 2[nd] theorems [2, p.15]. Indeed, by using (2.2) and the Merten's 1-st theorem (30), we have:

$$P\{\xi(n)=1\} = \prod_{p\le\sqrt{n}}\left(1-\frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\frac{1}{2}\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right] = \frac{c}{\ln(n)}\left[1+O\left(\frac{1}{\ln(n)}\right)\right] \tag{7.5}$$

where $c = \frac{2}{e^{\gamma}} \approx 1.122918968$. Setting $\lambda_n = \frac{1}{\ln n}$, we have $P\{\eta(n)=1\} = E\eta(n) \approx c\cdot\lambda_n$.

If we denote

$P\{\gamma_m(\nu)=1 \mid \nu = n\} = \beta(m,n)$, where $3 \le \nu \le 2m-3$, then, due to (7.4),

$$\beta(m,n) = \frac{1}{\ln(n)} \cdot \frac{1}{\ln(2m-n)}.$$

Since $G(2m,v_m) = \sum_{v_m=3}^{m-3} \gamma_m(v_m)$, we have for a mathematical expectation and a variance:

$$E\left\{G(2m,v_m)\right\} = \sum_{n=3}^{m-3} E\left\{\gamma_m(v_{m,n})\right\} = \sum_{n=3}^{2\cdot m-3} \beta(m,n) \sim \int_3^{2m-3} \beta(m,t)dt \qquad (7.6)$$

$$Var\left\{G(2m,v_m)\right\} = \sum_{n=3}^{m-3} Var\left\{\gamma_m(v_m)\right\}$$
$$= \sum_{n=3}^{2\cdot m-3} \left[\beta(m,n)\cdot(1-\beta(m,n)\right] \sim \int_3^{2m-3} \beta(m,t)\cdot\left(1-\beta(m,t)\right)dt$$

The main results of this section are stated in the following theorems.
The most critical question for the Goldbach strong Conjecture is whether the probability that for 'sufficiently large' values of $m > M \geq 3$ all sets $G_m\mathbb{P}$ are not empty, or equivalently, is this true that

$$P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,v_m) = |G_m\mathbb{P}| \;\; > 0\right\}\right\} \to 1 \text{ as } M \to \infty.$$

Let $G_m\mathbb{P}$ for $m \geq 3$ be a set of all G-primes, that is prime numbers $p, p' \in \mathbb{P}$ such that $p + p' = 2m$. Let each random variable v_m in the sequence of independent random variables $\left\{v_m\right\}_{3\leq m}$ follow Zeta probability distribution: $P\left\{v_m = n\right\} = \frac{n^{-s}}{\zeta(s)}$ $(s > 1)$ and $\gamma_m(v_m) = \begin{cases} 1 \text{ if } v_m = n \in G_m\mathbb{P} \\ 0 \text{ otherwise} \end{cases}$. Then $\left\{\gamma_m(v_m)\right\}_{m\geq 3}$ is a sequence of independent Bernoulli variables and the Goldbach function $G(2m,v_m) = \sum_{n=3}^{2m-3} \gamma_m(v_m)$ has the following properties:

(1) $$P\left\{G(2m,\nu_m)=0\right\}=P\left\{\bigcap_{n=3}^{2m-3}\left\{\gamma_m(\nu_m)=0\,|\,\nu_m=n\right\}\right\}\to 0 \text{ as } m\to\infty.$$

(2) $$\sum_{m=3}^{\infty}P\left\{G(2m,\nu_m)=0\right\}<\sum_{m=3}^{\infty}e^{-\frac{2m-6}{\ln^2(2m)}}\approx 6.00236$$

(3) $$\lim_{M\to\infty}P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,\nu_m)=|G_m\mathbb{P}|\ >0\right\}\right\}\to 1$$

Due to Lemma 3.2, the sequence $(\xi_k)_{k\in\mathbb{N}}$ in the Cramér's model is *asymptotically pairwise Bernoullian,* so that

$$\left|V\left\{X_n\right\}-\frac{1}{n^2}\sum_{k=1}^{n}V\left\{\xi_k\right\}\right|=O\left(\frac{1}{\ln N}\right) \text{ as } D_n=O\left(\frac{1}{\ln N}\right) \text{ for all } n>N.$$

In the formula $G(2m,\nu)=\sum_{n=3}^{2m-3}\gamma_m(n)>0$ we have all random variables $\gamma_m(n)$, $\mathrm{N}\le n\le 2m-3$, not necessarily independent but asymptotically linearly uncorrelated. Then, $\lim_{m\to\infty}P\left\{G_m\mathbb{P}\ge 1\right\}=1$.

Examples of sets $G_m\mathbb{P}$ for $2m=10,10^2,10^3$ with the corresponding values of $G(2m)$ are represented in the following table.

Table 7.1

G - primes in sets $G_m P$ and Goldbach function values

$2m$	$G_m\mathbb{P}$ sets	$G(2m)$
10	3 5 7	3
100	3 11 17 29 41 47 53 59 71 83 89 97	12
1000	3 17 23 29 47 53 59 71 89 113	
	137 173 179 191 227 239 257 281 317	
	347 353 359 383 401 431 443 479 491	
	509 521 557 569 599 617 641 647 653	
	683 719 743 761 773 809 821 827 863	
	887 911 929 941 947 953 971 977 983	
	997	56

In the context of the Goldbach conjecture we are interested in evaluation of the number of representations $G(2m)$ for an even number $2m$ in the form $2m = p + p'$, where $p, p' \in \mathbb{P}$. To incorporate a probabilistic approach in this context, let us consider a choice of a G-prime p for every $m \geq 3$ as a realization of a random variable $\nu = \nu_m$ with a certain probability distribution. We have then, $2m = \nu_m + \nu'_m$ where $\nu_m \in \mathbb{P}$, $\nu'_m \in \mathbb{P}$. Denote such a number $G(2m)$ as a realization of $G(2m, \nu_m)$. Since $G(2m,\nu_m) = \sum_{\nu_m=3}^{m-3} \gamma_m(\nu_m)$, we have for a mathematical expectation and a variance:

$$E\left\{G(2m,\nu_m)\right\} = \sum_{n=3}^{m-3} E\left\{\gamma_m(\nu_m)\right\} = \sum_{n=3}^{2\cdot m-3} \beta(m,n) \sim \int_3^{2m-3} \beta(m,t)dt \tag{7.7}$$

$$Var\left\{G(2m,\nu_m)\right\} = \sum_{n=3}^{m-3} Var\left\{\gamma_m(\nu_m)\right\}$$

$$= \sum_{n=3}^{2\cdot m-3} \left[\beta(m,n)\cdot(1-\beta(m,n)\right] \sim \int_3^{2m-3} \beta(m,t)\cdot\left(1-\beta(m,t)\right)dt$$

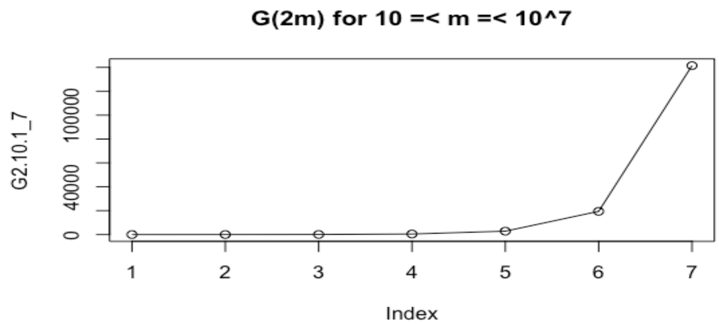


Figure 3

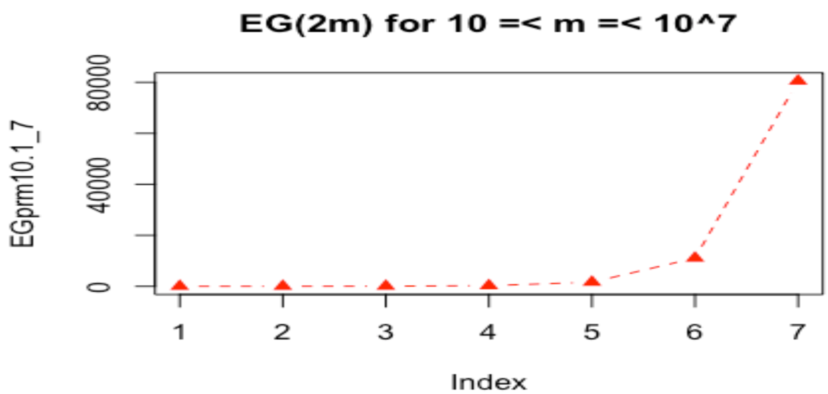


Figure 4

The Goldbach Conjecture can be stated in the form $G(2m,v_m)=\sum_{v_m=3}^{2m-3}\gamma_m(v_m)>0$ for all $m\geq 3$. Assumption that $G(2m,v_m)=0$ for some arbitrary large value of m contradicts to the increasing behavior of $G(2m,v_m)$when m increases. Moreover, we prove the following proposition.

Theorem 7.1

Let $G_m\mathbb{P}$ for $m\geq 3$ be a set of all G-primes, that is prime numbers $p,p'\in\mathbb{P}$ such that $p+p'=2m$. Let each random variable v_m in the sequence of independent random variables $\{v_m\}_{3\le m}$ follow Zeta probability distribution: $P\{v_m = n\} = \dfrac{n^{-s}}{\zeta(s)}$ $(s>1)$ and

$\gamma_m(v_m) = \begin{cases} 1 \text{ if } v_m = n \in G_m\mathbb{P} \\ 0 \text{ otherwise} \end{cases}$. Then $\{\gamma_m(v_m)\}_{m\ge3}$ is a sequence of independent

Bernoulli variables and the Goldbach function $G(2m,v_m) = \sum_{n=3}^{2m-3} \gamma_m(v_m)$ has the

following properties:

(1) $P\{G(2m,v_m)=0\} = P\left\{\bigcap_{n=3}^{2m-3}\{\gamma_m(v_m)=0 \mid v_m = n\}\right\} \to 0$ as $m\to\infty$.

(2) $\sum_{m=3}^{\infty} P\{G(2m,v_m)=0\} < \sum_{m=3}^{\infty} e^{-\frac{2m-6}{\ln^2(2m)}} \approx 6.00236$

(3) $\lim_{M\to\infty} P\left\{\bigcap_{m=M}^{\infty}\{G(2m,v_m) = |G_m\mathbb{P}| > 0\}\right\} \to 1$

Proof.

Independence of the Bernoulli variables in the set $\{\gamma_m(v_m) \mid v_m = n, 3\le n\le 2m-3\}$

follows from the assumed independence of v_m in the sequence $\{v_m\}_{3\le m}$ and

Theorem 4.1 regarding v_m with Zeta distribution. Then, due to independence of

$\{\gamma_m(v_m) \mid 3\le n\le 2m-3\}$, we have

$$P\left\{\bigcap_{n=3}^{2m-3}\{\gamma_m(v_m)=0 \mid v_m=n\}\right\} = \prod_{n=3}^{2m-3} P\{\gamma_m(v_m)=0 \mid v_m=n\} = \prod_{n=3}^{2m-3}\left[1-\beta(m,n)\right],$$

where $\beta(m,n) > \dfrac{1}{\ln^2(2m)}$. This implies:

$$P\{G(2m,v_m)=0\} < \prod_{n=3}^{2m-3}\left[1-\frac{1}{\ln^2(2m)}\right] = \left[1-\frac{1}{\ln^2(2m)}\right]^{2m-6} \sim e^{-\frac{2m-6}{\ln^2(2m)}} \to 0 \text{ as } m\to\infty.$$

We are interested in proving that $P\{G(2m,v_m)>0\}\to 1$ as $m\to\infty$.

A more critical question for the Goldbach Conjecture can be stated as follows: what is the probability that for 'sufficiently large' values of $m > M \geq 3$ all sets $G_m\mathbb{P}$ are not empty, that is

$$P\left\{\bigcap_{m=M}^{\infty}\left\{G(2m,v_m) = |G_m\mathbb{P}| \ > 0\right\}\right\} \to 1 \text{ as } M \to \infty.$$

Consider the probability of the opposite event: $P\left\{\bigcup_{m=M}^{\infty}\left\{G(2m,v_m)\right\} = 0\right\}$.

We have: $P\left\{\bigcup_{m=3}^{\infty}\left\{G(2m,v_m) = 0\right\}\right\} \leq \sum_{m=3}^{\infty} P\left\{G(2m,v_m) = 0\right\} < \sum_{m=3}^{\infty} e^{-\frac{2m-6}{\ln^2(2m)}} \approx 6.00236$.

Then, $P\left\{\bigcup_{m=M}^{\infty}\left\{G(2m,v_m)\right\} = 0\right\} \leq \sum_{m=M}^{\infty} P\left\{G(2m,v_m) = 0\right\} \to 0$ as $M \to \infty$

due to convergence of the series $\sum_{m=3}^{\infty} P\left\{G(2m,v_m) = 0\right\}$.

Q.E.D.

There is another way to evaluate the probability $P\left\lfloor G_m\mathbb{P} < 1\right\rfloor$.

Theorem 7.2

If in the formula $G(2m,v) = \sum_{n=3}^{2m-3} \gamma_m(n) > 0$ we assume (less restrictively) that all random variables $\gamma_m(n)$, $3 \leq n \leq 2m$-3, are not necessarily independent, but at least linearly uncorrelated. Then, $\lim_{m\to\infty} P\left\{G_m\mathbb{P} \geq 1\right\} = 1$.

Proof.

By applying the Central Limit Theorem, we have, due to formulas (1. 15), for sufficiently large values of m:

$$P\left\{G_m\mathbb{P} < 1\right\} = P\left\{X < x_{cr}(m)\right\} \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_{cr}(m)} e^{-\frac{1}{2}t^2}\, dt, \text{ where } x_{cr}(m) = \frac{1 - E\left\{G_m\mathbb{P}\right\}}{\sqrt{Var\left\{G_m\mathbb{P}\right\}}}.$$

Obviously, $\lim_{m\to\infty} P\{G_m\mathbb{P} < 1\} = 0$, since $x_{cr}(m) = \dfrac{1 - E\{G_m\mathbb{P}\}}{\sqrt{Var\{G_m\mathbb{P}\}}} \to -\infty$ as $m \to \infty$, which means that $\lim_{m\to\infty} P\{G_m\mathbb{P} \geq 1\} = 1$.

Q.E.D.

The values of $P\{G_m\mathbb{P} < 1\}$ and $x_{cr}(m)$ for $m = 10^3, 10^4, \ldots, 10^8$ are given in the following table.

m	10^3	10^4	10^5	10^6	10^7	10^8
$x_{cr}(m)$	-6.866973	-16.130926	-40.343498	-105.469447	-284.348502	-783.836910
$P\{G(2m) < 1\}$	3.278916×10^{-12}	7.734173×10^{-59}	0.0000000	0.0000000	0.0000000	0.0000000

8. Diffusion Approximations for Counting Number of Primes

In this chapter we consider the sequence $\{\pi(n)\}_{n\in\mathbb{N}}$ as a realization of a random walks $\{\pi(n,\omega)\}_{n\in\mathbb{W}}$ generated by the recurrent equation

$$\pi(n_{k+1})-\pi(n_k)=\eta(n_{k+1}) \text{ where } \eta(n_k)=h\big(\min(\vec{r}(n_k))\big),\ n_k=\nu_k(\omega) \tag{8.1}$$

Here $\{\nu_k\}_{k\in\mathbb{N}}$ are assumed to be random variables with Zeta probability distribution. Recall that to define a stochastic process $\xi(t,\omega)$ with a discrete or a continuous set $\mathcal{X}$ of values we need to have a measurable space $(\mathcal{X},\mathcal{B})$, where often $\mathcal{X}\subseteq\mathbb{N}$ or $\mathcal{X}\subseteq\mathbb{R}^d$, a Borel σ-algebra $\mathcal{B}$ of subsets on $\mathcal{X}$, and a set T of parameters $t\in T$ such that for each $t\in T$, $\xi(t,\cdot):\Omega\to\mathcal{X}$ is a random variable on a probability space $(\Omega,\mathcal{F},P)$. Then, the family $\{\xi(t,\cdot)\}_{t\in T}$ of random variables is called a *stochastic process* in the *phase space* $(\mathcal{X},\mathcal{B})$. The parameter $t\in T$ is usually interpreted as 'discrete time' for a countable set $T\subseteq\mathbb{W}=\mathbb{N}\cup\{0\}$ or as 'continuous time' for the continuous interval $T=\big[t_0,t_f\big)\subseteq\mathbb{R}^+=\big[0,\infty\big)$.

Then, a $\xi(t,\omega)$ is called a *stochastic process* with a discrete or a continuous time, respectively. For any given elementary event $\omega\in\Omega$, a function $x(\cdot):T\to\mathcal{X}$ such that $x(t)=\xi(t,\omega)$ is called a *path* (or a *trajectory*) of the random process. Alternatively, a stochastic process can be defined as a collection of paths (*random elements)* $x(\cdot)=\xi(\cdot,\omega)$ in a function space $\mathcal{X}^T=\{x(t)\,|\,t\in T\}$ where ω that identifies each path is an elementary event in probability space $(\Omega,\mathcal{F},P)$. Elements (or points) $x\in\mathcal{X}$ are called '*states*' of the process, and $\xi(t,\omega)$ itself is called a process with a discrete or continuous phase space $(\mathcal{X},\mathcal{B})$.

Following the historical traditions of the classical Probability Theory (and the development of Calculus, in general), we try to apply limit theorem approach to analyze behavior of infinite discrete random sequences in terms of continuous-time stochastic processes.

Let $\vec{p}(n) = (p_1, p_2, \ldots, p_k)$ be a vector of consecutive prime numbers such that $p_1 = 2,\ p_k \le n$ and $p_{k+1} > n$. Index k determines here the value of function $\pi(n) = k$ that is the number of primes less than or equal to n so that $\vec{p}(n) = (p_1, p_2, \ldots, p_{\pi(n)})$. For each coordinate p_i of vector $\vec{p}(n)$ we determine the residual value $r_i = \mathrm{mod}(n, p_i),\ i = 1, 2, \ldots, \pi(n),$ and consider the corresponding vector of residuals $(r_1, r_2, \ldots, r_{\pi(n)})$. Notice that, due to the Sieve Algorithm, for an integer $n > 2$ to be prime it is necessary and sufficient that the all coordinates $r_i \ \left(1 \le i \le \sqrt{\pi(n)}\right)$ of the 'reduced' vector of residuals $\vec{r}(n) = \left(r_1, r_2, \ldots r_{\sqrt{\pi(n)}}\right)$ be different from zero (Lemma 1.2). Meanwhile, if a random integer ν follows Zeta distribution, then, due to Lemma 2.1, the events that $\nu = n$ does not divide each of consecutive primes $p_1, p_2, \ldots, p_{\pi(\sqrt{n})}$ are independent and can be expressed as a condition:

$$\min\left\{r_i \mid 1 \le i \le \pi\left(\sqrt{n}\right)\right\} > 0 \text{ or, equivalently, } \prod_{1 \le i \le \pi(\sqrt{n})} r_i > 0. \qquad (8.2)$$

By using the Heaviside function $h(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}$, we can write the recurrent equation for $\pi(n)$ in the form:

$$\pi(n+1) = \pi(n) + h\left(\min_{p \le \sqrt{n}}\left\{\mathrm{mod}(n, p) \mid p \in \mathbb{P}\right\}\right)$$

or, equivalently,

$$\pi(n+1) = \pi(n) + h\left(\min_{i \le \sqrt{n+1}}\left\{r_i \mid r_i = \mathrm{mod}(n, p_i)\right\}\right) = \pi(n) + h\left(\min(\vec{r}(n)\right) \qquad (8.3)$$

which controls the occurrence of prime numbers in the sequence of all integers $n \geq 3$: $h\left(\min(\vec{r}(n)\right) = 1$ if and only if n is a prime number and $h\left(\min(\vec{r}(n)\right) = 0$ otherwise.

Consider a stochastic process approximation of non-Markov random walks $\left\{\pi(n,\omega)\right\}_{n\in\mathbb{W}}$ such that $\pi(n,\omega) = \pi\left(n\right)$, with $\pi : \mathbb{N}\times\Omega \to \mathbb{N}\cup\left\{0\right\}$ restricted to the interval of discrete 'times' $N_{\min} = n_0 < n_1 < \cdots < n_K = N_{\max}$:

$$\left\{\pi^{\Delta}(t_k) = \pi(n_k,\omega) \mid N_{\min} \leq n_k \leq N_{\max}\right\}. \tag{8.4}$$

Denote $\Delta = \left(0 = t_0 < t_1 < \ldots < t_K = 1\right)$ a partition of an interval $\left[0,1\right]$ into K subintervals, such that $\dfrac{K}{\ln\left(N_{\max}\right)} \to 0 \ as \ N_{\max} \to \infty$.

We can map the closed interval of real numbers $\left[N_{\min}, N_{\max}\right] \subset \mathbb{R}$ to the interval $\left[0,1\right] \subset \mathbb{R}$ with an increasing continuously differentiable function $\tau(x)$ such that $\tau(N_{\min}) = 0,\ \tau(N_{\max}) = 1$.

In the context of our study, a suitable choice of function τ takes the form:

$$\tau(x) = \frac{\int\limits_{N_{\min}}^{x} \frac{dt}{\ln t}}{\int\limits_{N_{\min}}^{N_{\max}} \frac{dt}{\ln t}} = \frac{Li(x) - Li(N_{\min})}{Li(N_{\max}) - Li(N_{\min})} \tag{8.5}$$

where $Li(x)$ stands for the *Eulerian logarithmic integral* $Li(x) = \int\limits_{2}^{x} \frac{dt}{\ln t}$.

Then, $t_k = \tau(n_k)$ and for τ^{-1} (the inverse of τ) we have $n_k = \tau^{-1}(t_k)$ $(k = 1,2,\ldots,K)$.

Denote $\Delta t_k = t_k - t_{k-1}$. Assume that $N_{\min} \to \infty$ and for each choice of $N_{\min}$ a positive integer K can be taken such that $\left|\Delta\right| = \max\limits_{1\leq k\leq K} \Delta t_k \to 0$. Here a sequence of random variables $\pi^{\Delta}(t_k) = \pi(n_k)$ is interpreted as a path of a walking point $\pi^{\Delta}(t_k)$ that

belongs to a measurable space $\left(\mathcal{X}_k,\mathcal{B}_k\right)$ at each 'instant of registration' t_k.
Probability distribution on the probability space $\left(\Omega,\mathcal{F},P\right)$ generated by the path space $\left(\mathcal{X}^\infty,\mathcal{F}^\infty\right)=\left(\prod_{k\in\mathbb{N}}\mathcal{X}_k,\underset{k\in\mathbb{N}}{\otimes}\mathcal{B}_k\right)$ of random walks $\left\{\pi(k,\omega)\right\}_{k\in\mathbb{W}}$ is determined by transition probabilities $P\left\{\pi^\Delta(t_{k+1})\in E\,|\,\pi^\Delta(\vec{t}_0^{\,k})=\vec{x}_0^{\,k}\right\}$ where $\vec{t}_0^{\,k}=\left(t_0,t_1,\ldots,t_k\right)$, $\vec{x}_0^{\,k}=\left(x_0,x_1,\ldots,x_k\right)\in\mathcal{X}_0^k=\prod_{i=0}^{k}\mathcal{X}_i$, $x_i\in\mathcal{X}_i\ (i=0,1,\ldots)$, $E\in\mathcal{B}_{k+1}$. Existence and uniqueness of the probability path space $\left(\Omega,\mathcal{F},P\right)$ follows from the theorem of Ionescu Tulcea [30].

Notice that $\pi:\mathbb{N}\to\mathbb{W}=\mathbb{N}\cup\left\{0\right\}$ and therefore, we set $\mathcal{X}_k=\mathbb{W}$ for all $k\in\mathbb{N}^*$.

To prove the weak convergence of transition probabilities for the sequence of random walks (3.1.4) to the diffusion process $\hat{\pi}(t)$ on the time interval $[0,1]$, consider so called *infinitesimal characteristics* of the random walks:

$$m^\Delta(t_k,\vec{x}_1^{\,k})=\frac{1}{\Delta t_k}E\left\{\Delta\pi(t_{k+1})\,|\,\pi^\Delta(\vec{t}_1^{\,k})=\vec{x}_1^{\,k}\right\},$$

$$\left[\sigma^\Delta(t_k,\vec{x}_1^{\,k})\right]^2=\frac{1}{2\cdot\Delta t_k}E\left\{\left[\Delta\pi(t_{k+1})\right]^2|\,\pi^\Delta(\vec{t}_1^{\,k})=\vec{x}_1^{\,k}\right\}$$

$$g^\Delta\left(t_k,\vec{x}_1^{\,k};\Gamma_{k+1}\right)=\frac{1}{\Delta t_k}E\left\{I_{\Gamma_{k+1}}(\Delta\pi^\Delta(t_{k+1})\,|\,\pi^\Delta(\vec{t}_1^{\,k})=\vec{x}_1^{\,k}\right\}.$$

Here $\Delta\pi(t_{k+1})=\pi^\Delta(t_{k+1})-\pi^\Delta(t_k)=\eta^\Delta(t_{k+1})=\pi(n_{k+1})-\pi(n_k)=\eta(n_{k+1})$

$\Delta\pi(t_{k+1})=\pi^\Delta(t_{k+1})-\pi^\Delta(t_k)=\pi(n_{k+1})-\pi(n_k)=\eta^\Delta(t_{k+1})=\eta(n_{k+1})$, $\vec{t}_1^{\,k}=\left(t_1,t_2,\ldots,t_k\right)$,

$\pi^\Delta(\vec{t}_1^{\,k})=\left(\pi^\Delta(t_1),\pi^\Delta(t_2),\ldots,\pi^\Delta(t_k)\right)=\left(x_1,x_2,\ldots,x_k\right)=\vec{x}_1^{\,k}$, $I_\Gamma(x)=\begin{cases}1 \text{ if } x\in\Gamma\\ 0 \text{ otherwise}\end{cases}$,

$\Gamma_{k+1}\subset\mathcal{X}_{k+1}\setminus\left\{x_k\right\}\in\mathcal{B}_{k+1}$.

By setting $n_k = n_0 + k$ for all $k = 0,1,\dots,K$, we have:

$$\Delta\pi^{\Delta}(t_{k+1}) = \pi(n_{k+1}) - \pi(n_k) = \eta(n_k + 1) = \eta^{\Delta}(t_{k+1})$$

$$\Delta\pi^{\Delta}(t_{k+1}) = \pi(n_{k+1}) - \pi(n_k) = \pi(n_k + 1) - \pi(n_k) = \eta(n_{k+1}) = \eta^{\Delta}(t_{k+1})$$

$\Delta\pi^{\Delta}(t_{k+1}) = \pi(n_{k+1}) - \pi(n_k) = \pi(n_k + 1) - \pi(n_k) = \eta(n_k + 1) = \eta^{\Delta}(t_{k+1})$, where

$$\begin{aligned} P\left\{\eta^{\Delta}(t_{k+1}) = 1\right\} &= r_{n_{k+1}} = \prod_{p \le \sqrt{n_{k+1}}} \left(1 - \frac{1}{p}\right), \\ P\left\{\eta^{\Delta}(t_{k+1}) = 0\right\} &= 1 - r_{n_{k+1}} \end{aligned} \tag{8.6}$$

We have then,

$$m^{\Delta}(t_k, \vec{x}_k) == \frac{1}{\Delta t_k} \cdot E\left\{\eta(n_{k+1})\right\} = \frac{1}{\Delta t_k} \cdot \prod_{p \le \sqrt{n_{k+1}}} \left(1 - \frac{1}{p}\right) \tag{8.7}$$

Similar, since $\eta(n_{k+1}) = \left[\eta(n_{k+1})\right]^2$, we have

$$\left[\sigma^{\Delta}(t_k, \vec{x}_k)\right]^2 = \frac{1}{\Delta t_k} \cdot E\left\{\left[\eta(n_{k+1})\right]^2\right\} = \frac{1}{\Delta t_k} \cdot \prod_{p \le \sqrt{n_{k+1}}} \left(1 - \frac{1}{p}\right) \tag{8.8}$$

By applying the first Merten's theorem to (6.5) and (6.6), we have

$$\begin{aligned} m^{\Delta}(t_k, x_{n_k}) &= \frac{1}{\Delta t_k} \cdot \frac{c}{\ln(n_k + 1)} \cdot \left[1 + O\left(\frac{1}{\ln(n_k + 1)}\right)\right] \\ \left[\sigma^{\Delta}(t_k, x_{n_k})\right]^2 &= \frac{1}{\Delta t_k} \cdot \frac{c}{\ln(n_k + 1)} \cdot \left[1 + O\left(\frac{1}{\ln(n_k + 1)}\right)\right] \end{aligned} \tag{8.9}$$

LEMMA 9.1

For any interval $[a,b]$ with integer a and b such that $0 < a < b$, we have

$$\left|\sum_{a < n \le b} \frac{1}{\ln n} - \int_a^b \frac{dt}{\ln t}\right| \le \int_a^b \frac{dt}{t \cdot (\ln t)^2} \le \frac{b - a}{a \cdot (\ln a)^2} \qquad (8.10))$$

Proof.

Due to the *Euler's summation formula* [12, p.54], for positive integer numbers a and b and a function f with a continuous derivative f' on $[a,b]$, we have

$\sum_{a<n\le b} f(n) = \int_a^b f(t)\,dt + \int_a^b \left(t-[t]\right) f'(t)\,dt$,where $[t]$ denotes an integer part of t .

$$\left|\sum_{a<n\le b}\frac{1}{\ln n} - \int_a^b \frac{dt}{\ln t}\right| = \left|\int_a^b \left(t-[t]\right) f'(t)\,dt\right| \le \int_a^b \left|f'(t)\right| dt \le \frac{b-a}{a\cdot(\ln a)^2}$$

Q.E.D.

Consider the *Eulerian logarithmic integral* $Li(x) = \int_2^x \frac{dt}{\ln t}$ to evaluate $\sum_{i=n_k+1}^{n_{k+1}} r_{ki}$.

LEMMA 9.2.

$$\sum_{i=n_k+1}^{n_{k+1}} r_{ki} = \int_{n_k}^{n_{k+1}} \frac{dt}{\ln t} = Li(n_{k+1}) - Li(n_k) + O\left(\frac{n_{k+1}}{\ln^2(n_{k+1})}\right) = \frac{n_{k+1}}{\ln n_{k+1}} - \frac{n_k}{\ln n_k} + O\left(\frac{n_{k+1}}{\ln^2 n_{k+1}}\right)$$

Proof.

We have

By using approximation [12]: $li(x) = \int_0^x \frac{dt}{\ln t} = \frac{x}{\ln x} + O\left(\frac{1}{\ln^2 x}\right)$,

we have: $Li(x) = li(x) - li(2),$ where $li(x) = \int_0^x \frac{dt}{\ln t} = \frac{x}{\ln x} + O\left(\frac{x}{\ln^2 x}\right).$

This implies $Li(n_{k+1}) - Li(n_k) = \frac{n_{k+1}}{\ln n_{k+1}} - \frac{n_k}{\ln n_k} + O\left(\frac{n_{k+1}}{\ln^2 n_{k+1}}\right)$.

Q.E.D.

Consider now a diffusion process $\hat{\pi}(t)$ given by the sum of stochastic integrals:

$$\hat{\pi}(t) = \int_0^t m(s)\,ds + \int_0^t \sigma(s)\,dw(s) \qquad (8.11)$$

where $m(t)=\dfrac{c}{\ln\left(\tau^{-1}(t)\right)}, \sigma(t)=\dfrac{1}{2}\cdot m(t)\cdot\left(1-m(t)\right), c=\dfrac{2}{e^{\gamma}}, 0\le t\le 1; \tau^{-1}(t)=x$.

with the transition probability $u(t,x,A)=P\left\{\hat{\pi}(t)\in A \mid \hat{\pi}(t_0)=0\right\}$.

Here $w(t)$ is a process of Brownian motion on $0\le t\le 1$.

The semigroup of linear operators U_t is defined on the space of bounded measurable functions by $\left(U_t f\right)(x)=\int f(y)u(t,x,dy)$.

We have the *infinitesimal generator of the semigroup* U_t given by the formula:

$$\left(Lf\right)(x)=\lim_{\Delta t\to 0}\frac{\left(U_{t+\Delta t}f\right)(x)-f(x)}{\Delta t} .$$

On the set of twice continuously differentiable functions $C^2(\mathbb{R})$ the generator L takes a form of a differential operator $\left(Lf\right)(x)=m(x)\dfrac{\partial f}{\partial x}+\dfrac{1}{2}\sigma^2(x)\dfrac{\partial^2 f}{\partial x^2}$.

The function $V(t,x)=\left(U_t f\right)(x)=E\left[Y(t)\mid Y(t_0)=x\right]=\int f(y)u(t,x,dy)$ satisfies the equation

$$\frac{\partial V}{\partial t}=\frac{1}{2}\sigma^2(x)\frac{\partial^2 V}{\partial x^2}+m(x)\frac{\partial V}{\partial x} \tag{8.12}$$

with the initial condition $V(t_0,x)=f(x)$. Taking as an initial condition δ-function, we have $V(t,x)=u(t,x,y)$, called a *fundamental solution* to (6.9).

This means that the transition probability has a density $u(t,x,y)$, so that

$$P\left\{Y(t)\in A \mid Y_0(t)=x\right\}=\int_A u(t,x,y)\,dy .$$

By applying the generalized limit theorem [31, 32] about convergence of random walks $\pi^{\Delta}\left(t_k\right)$ as $\left|\Delta\right|=\max_{1\le k\le K}\Delta t_k\to 0, N_{\min}\to\infty$ to diffusion processes (6.9), we obtain an approximation of $\left\{\pi(n,\omega)\right\}_{n\in\mathbb{W}}$ in terms of diffusion processes, defined for expanding intervals $[N_{\min},N_{\max}]$ of approximation on $\mathbb{N}$.

Theorem 9.1

Transition probabilities

$P\left\{\pi^{\Delta}(t_{k+1}) \in E \mid \pi^{\Delta}(t_k) = x_k, \pi^{\Delta}(t_{k-1}) = x_{k-1}, \ldots, \pi^{\Delta}(t_0) = x_0\right\}$, where $\vec{x}_k = (x_1, \ldots, x_k) \in \mathbb{N}^k$,

of the defined above non-Markov random walks $\left\{\pi^{\Delta}(t_k) \mid N_{\min} < k < N_{\max}\right\}$ *converge weakly* to the transition probabilities of the diffusion process $\hat{\pi}(t)$ given by the stochastic integral

$$\hat{\pi}(t) = \int_0^t \hat{m}(s)\, ds + \int_0^t \hat{\sigma}(s)\, dw(s),$$

where $\hat{m}(t) = \dfrac{c}{\ln\left(\tau^{-1}(t)\right)}, \hat{\sigma}(t) = \dfrac{1}{2} \cdot \hat{m}(t) \cdot \left(1 - \hat{m}(t)\right), c = \dfrac{2}{e^{\gamma}}, 0 \le t \le 1,$

$\tau^{-1}(t) = x,\ N_{\min} \le n \le N_{\max}, \tau(N_{\min}) = 0, \tau(N_{\max}) = 1,\ c = \dfrac{2}{e^{\gamma}} \approx 1.122918968$

with the Euler's constant $\gamma = \sum\limits_{m \le n} \dfrac{1}{m} - \ln n + O\left(\dfrac{1}{n}\right), \gamma \approx 0.577215664$,

as $\left|\Delta\right| = \max\limits_{1 \le k \le K} \Delta t_k \to 0,\ N_{\min} \to \infty$.

Proof.

Since $\sum\limits_{k=1}^{K} \Delta t_k = 1$, to due to Lemma 3.1.1, we have:

$$\sum_{k=1}^{K} \frac{1}{\ln n_k} \le \frac{K}{\ln(N_{\min})} \cdot \left(1 + O\left(\frac{1}{\ln(N_{\min})}\right)\right) \to 0, \text{ while } \frac{K}{N_{\max}} \to 0$$

Then, formulas (3.1.7), due to the second Merten's theorem (the *Merten's formula*) [2, p.19], imply:

$$\sum_{k=1}^{K}\left[\,|\,m^{\Delta}(t_k,\vec{x}_k)-m(t_k)\,|+|\left(\sigma^{\Delta}(t_k,\vec{x}_k)\right)^2-\left(\sigma(t_k,\vec{x}_k)\right)^2\,|\,\right]\cdot\Delta t_k$$

$$=2\cdot\sum_{k=1}^{K}\left[\left|\frac{1}{\ln(n_k)}-\frac{1}{\ln(n_k)}\cdot\Delta t_k+\frac{1}{\ln(n_k)}\cdot O\left(\frac{1}{\ln(n_k)}\right)\right|\right]\le 2\cdot\sum_{k=1}^{K}\left[\frac{1}{\ln(n_k)}\left(1-\Delta t_k+O\left(\frac{1}{\ln(n_k)}\right)\right)\right]$$

$$\le 2\cdot\max_{1\le k\le K}\left|1-\Delta t_k+O\left(\frac{1}{\ln(n_k)}\right)\right|\cdot\sum_{k=1}^{K}\frac{1}{\ln n_k}\to 0,\ \text{as } \mathrm{N}_{\min}\to\infty$$

For $g(k)=\pi(k+1)-\pi(k)\le 1$, we have $P\left\{\eta^{\Delta}(k)>1\right\}=0$ for all k, so that all conditions are satisfied to apply the limit theorems for random walks proved in [31 ,32, 33].

Q.E.D.

The figures below illustrate graphically the diffusion approximation of distribution of primes in terms of $\pi(n)$ on different intervals of the argument.

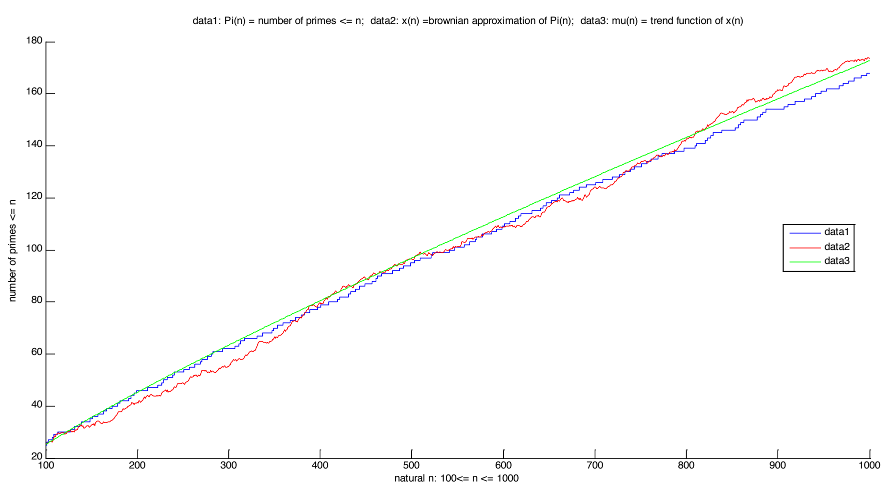
Legend for the graphs on the following figures:

data1: $\pi(n)=$ exact number of primes $\le$ n

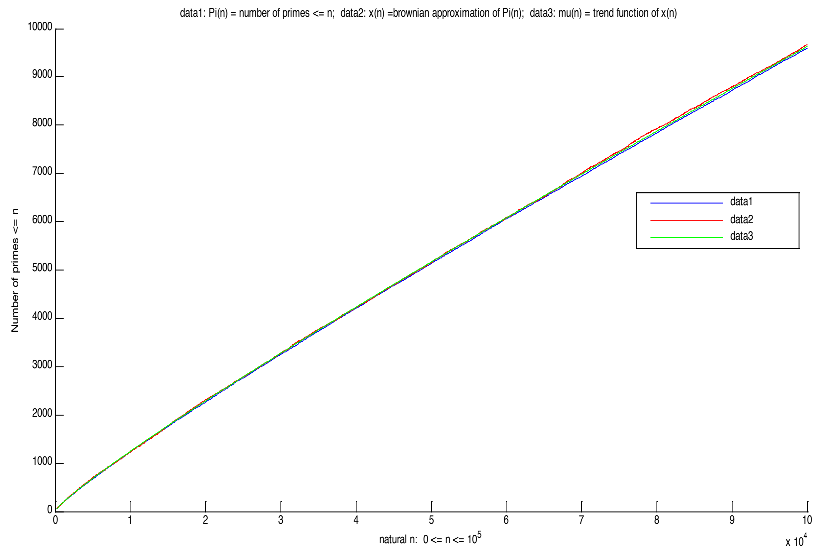
data2: Brownian approximation $X_n=\mu(n)+\xi_n\cdot\sigma(n)$ of $\pi(n)$

data 3: Trend function $\mu(n)$ of X_n

Approximation of $\pi(n)$ for $n:100\le n\le 1000$



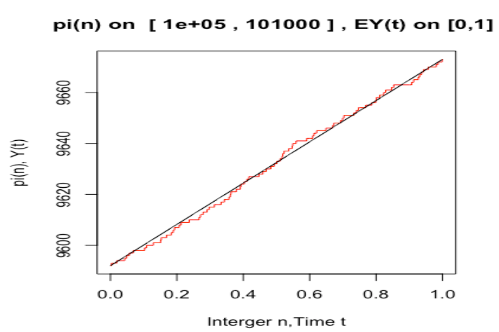
Approximation of $\pi(n)$ for $n: 0 \le n \le 10^5$

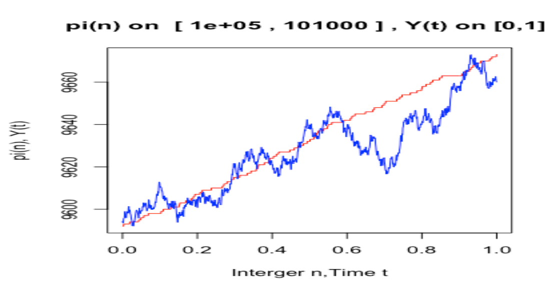


Approximation of $\pi(n)$ for $n: 5\times 10^5 \le n \le 10^6$

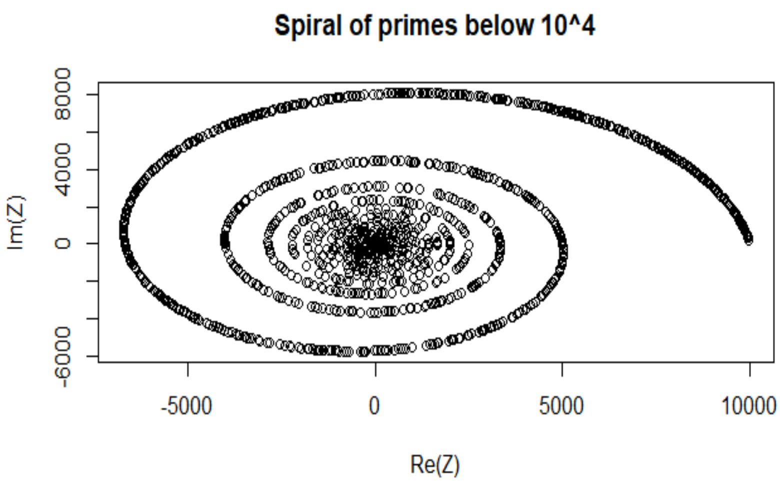
On figures below there are the graphs of paths described evolution of the 'walk' of a counts $\{\pi(n) \mid n \in \mathbb{N}\}$ of consecutive primes restricted to the intervals $[N_{\min}, N_{\max}] \subset \mathbb{N}$ and approximating diffusion processes :

$Y(t) = \hat{\pi}^{\Delta}(t)$ and their expectations $\mathrm{E}Y(t)$ for $t \in [0,1] \subset \mathbb{R}$.

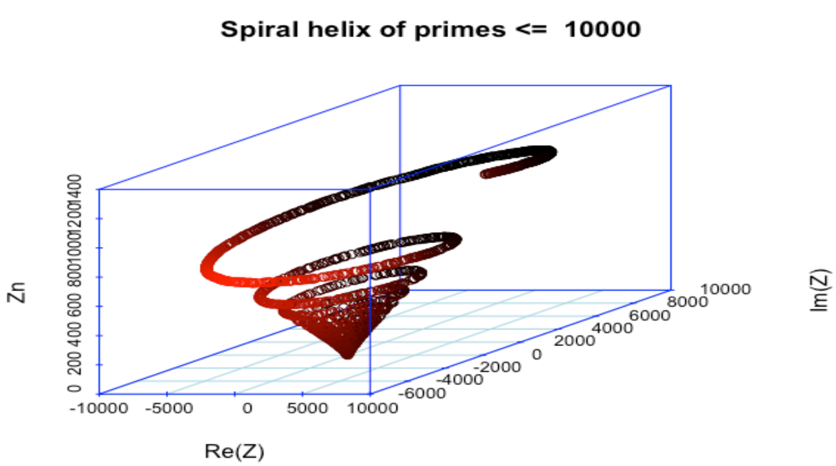




The sequence of vectors $(\vec{p}(n), \vec{r}(n))$, $(n = 2, 3, \ldots)$ created by consecutive n primes and the residual values $\vec{r} = \mathrm{mod}(n, \vec{p})$, allows an interesting 3D presentation. In each pair $(\vec{p}(n), \vec{r}(n))$ vector of primes $\vec{p}(n)$ represents a 'radial' component, while the vector of residuals $\vec{r}(n)$, due to its natural periodicity, represents a 'circular' component.



Denote $z_k = p_k \cdot \exp\left(2\pi i \cdot \frac{r_k}{p_k - 1}\right)$, $r_k = \mathrm{mod}(n, p_k)$, $(k = 1, 2, 3, \ldots)$ - a sequence of complex numbers and the vector $\vec{z}(n) = (z_1, \ldots, z_n)$. Then for any $n > 2$ vector $\vec{z}(n)$ takes a shape of a spiral helix as in the pictures below.



About the author

Gregory M. Sobko is a retired professor of mathematics at National University, San Diego, and at UC San Diego Extension. He obtained his MS degree in 1967 and PhD in Mathematics in 1973 from Moscow State University, Russia. His works are in Probability Theory (Limit Theorems for random walks on Lie groups and differentiable manifolds) and applications (Random Fields, Stochastic Differential Equations).

REFERENCES

1. Yuang Wang, *The Goldbach Conjecture.* Second Edition. World Scientific. Singapore, London, 2002.
2. Cramér Harald, *On the order of magnitude of the difference between consecutive prime numbers,* Acta Arithmetica, 2: 23-46.1936.
3. Granville, A., *Harald Cramér and the distribution of prime numbers*, Scandinavian Actuarial Journal, 1:12-28. 1995]
4. Mark Kac, *Statistical Independence in Probability, Analysis and Number Theory*, The Mathematical Association of America, John Wiley and Sons, Inc.,1972.
5. Tom M. Apostol, *Introduction to Analytic Number Theory*, Springer, 2000.
6. Edward Beltrami, *What is Random? Chance and Order in Mathematics and Life.* Springer-Verlag New York, Inc. , 1999.
7. Song Y. Yan, *Number Theory for Computing*, Springer, Springer-Verlag, 2000.
8. A.N. Shiryaev, *Probability*, 2nd edition, Springer, 1996.
9. V.S. Varadarajan, *Euler Trough Time: A new Look at Old Themes*, AMS, 2006.
9. Gérald Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, 3rd Edition, AMS, 2015.
11. Gérald Tenenbaum and Michael Mendès , *The Prime Numbers and Their Distribution,* AMS, 2000.
12. А. Г. Постников, *Введение в аналитическую теорию чисел*, Москва, Наука, 1971.
13. Gwo Dong Lin and Chin-Yuan Hu, *The Riemann zeta distribution*, Bernoulli 7(5), 2001, pp. 817-828.

14. Philippe Biane, Jim Pitman, and Mark Yor, *Probability Laws related to the Jacobi Theta and Riemann Zeta Functions, and Brownian Excursions*, Bulletin o the American Mathematical Society. Vol. 38, Number 4, p. 435-465, 2001.
15. Solomon W. Golomb, *A Class of Probability Distributions on the Integers*. Journal of Number Theory 2, 1970, p. 189-192.
16. Andrew Granville, *Primes in Intervals of Bounded Length*, Bulletin of the American Mathematical Society, Vol.52, Number 2, April 2015, p.171-222.
17. Bryan Kra, *Poincaré Reccurence and Number Theory: Thirty Years Later*. Bulletin of the American Mathematical Society, Vol.48, Number 4 October 2011, p. 497–501.
18. Harry Fusrstenberg, *Poincaré Reccurence and Number Theory*, Bulletin of the American Mathematical Society, Volume 5, Number 3, November 1981, p. 211-234.
19. Ludwig Arnold, *Random Dynamical Systems*, Springer-Verlag, Berlin Heidelberg, 1998
20. V.K. Balakrishnan, Introductory Discrete Mathematics, Prentice-Hall International, Inc.,1991.
21. Michel Loèv, Probability Theory, Third Edition. Dover Publications, Inc., Mineola, New York, 2017.
22. G.M. Sobko, *The First Diffusion Problem on Differentiable Manifolds*, Theory of Probability and its Applications, v. XVII, n.3, 1972, pp. 521 – 528.
23. G. M. Sobko, *A Diffusion Approximation of non-Markov Random Walks on Differentiable Manifolds*, Theory of Probability and its Applications, v. XVIII, n.1, 1973, pp. 41 - 53

24. Г.М. Собко, *Предельные теоремы для случайных блужданий на дифференцируемых многообразиях*. Механико-математический факультет, Механико-математический факультет, МГУ, 1973. *(*G. M. Sobko*, Limit Theorems for Random Walks on Differentiable Manifolds*. Dissertation. Department of Mathematics and Mechanics. Moscow State University, 1973,pp. 1-148.

25.J. Aczél, *Lectures on Functional Equations and Their Applications*, Vol.14, Academic Press, New York and London, 1966.

26. K. K. Mardzanishvili and A.V. Postnikov, *Prime Numbers* (p. 199 – p. 228) in D. Aleksandrov, A. N. Kolmogorov, and M.A. Lavrent'ev. Mathematics, Its Contents, Methods and Meaning. Dover Publications, Inc., Mineola. New York, 1999.

27. А. Г. Постников, *Введение в аналитическую теорию чисел*. Москва, "Наука", 1971.

28. А. Н. Колмогоров, *Основные понятия теории вероятностей*. Москва, "Наука", 1974.

29. A.N. Shiryaev, *Probability*, 2nd edition, Springer, 1996.

30. S.R.S Varadhan, *Probability Theory*. Courant Institute of Mathematical Studies, American Mathematical Society. 2001.

31. S.R.S Varadhan, *Stochastic Processes*. Courant Institute of Mathematical Studies, American Mathematical Society. 2007.

32. Yu. I. Manin, Alexey A. Panchishkin, *Introduction to Modern Number Theory*. Springer, 2006

33. Wang Liang, Huang Yan, *A kind of dynamic model of prime numbers*. https://arxiv.org/pdf/math/0601517 , Department of Control Science and Control Engineering, Huazhong University of Science and Technology.

34. Radoslaw Hofman, *Why LP cannot solve large instances of NP-complete problems in polynomial time*. Poznan, 2006
35. Richard K. Guy, *Unsolved Problems in Number Theory. Second Edition*. Springer-Verlag, New York, 1994.
36. Yuang Wang, *The Golbach Conjecture*. Second Edition, World Scientific. Singapore, London, 2002.
37. Patrick Billingsley, *Prime Numbers and Brownian Motion*. The American Mathematical Monthly, Vol.80, No. 10 (Dec. 1973), pp.1099-1115.
38. Mark B. Villarino Depto, *Mertens' Proof of Mertens' Theorem. Depto de Matema'tica, Universidad de Costa Rica, 2060 San Jose',* Costa Rica April 28, 2005
39. J. Galambos. *Distribution of arithmetic functions. A survey*. Annales de l'I. H.P., section B, tome 6, n. 4(1970), p. 281-305.
40. Marek Wolf, *Random walk on the prime numbers*, Physica A 250 (1998). Institute of Theoretical Physics, University of Wroclaw, Poland.
41. Jacque Neveu, *Bases Mathématiques du Calcul des Probabilités*, Masson et Cie, Paris, 1964.

2010 Mathematics Subject Classification
Primary 11Nxx, 11N05, 11A41,11A07, 11A25, 11A51,
11N60, 11N37, 11P32, 11Mxx, 11M06, 11Y05;
Secondary: 11N36, 11Y16, 60-xx, 60Fxx, 60J65, 60G50,
60Bxx, 60B15

email: gbeholder@gmail.com (Gregory Sobko)
tel. (858)705-4143